Master thesis in Financial Economics

# Concentrated wealth creation in US industries: a few winners take all

*Author:*
Sam Nico Brasser

*Supervisor:*
Dr. Jan Lemmen

*Student ID:*
508065sb

*Co-reader:*
Dr. Rogier Quaedvlieg

**Abstract**

Severe asymmetry in long horizon stock returns and concentration of wealth creation can be detrimental for investors holding a concentrated portfolio of stocks. Therefore, this paper investigates the distribution of long horizon stock returns, concentration in wealth creation and the ability of stocks to outperform the risk-free asset in US industries from 1926 to 2018. We find that stock returns are extremely and increasingly skewed over longer horizons, the majority of stocks in US industries is unable to outperform the risk-free rate or inflation and that 23 out of 50 industries have a negative median lifetime buy-and-hold return. In the most concentrated industries less than 1% of firms is responsible for all net wealth creation. Large differences also occur in absolute wealth creation, with some industries creating close to zero wealth in the entire period from 1926 to 2018 while others create trillions of dollars in wealth. Fractional regressions are used to explain the cross-section of concentration in wealth creation in US industries. This empirical research finds that short-term return characteristics, corporate performance and variability in corporate performance can explain variation in cross-industry concentration of wealth creation and performance against the risk-free asset.

**JEL classification:** G11; G12; G32

**Keywords:** long-term returns; return benchmarking; compounding; return skewness; corporate performance; fractional regressions; concentration of wealth creation

Rotterdam, April 2020

# Contents

# List of Tables

# List of Figures

## Non-plagiarism statement

By submitting this thesis the author declares to have written this thesis completely by himself, and not to have used sources or resources other than the ones mentioned. All sources used, quotes and citations that were literally taken from publications, or that were in close accordance with the meaning of those publications, are indicated as such.

## Copyright statement

# 1 Introduction

Long-term investors could be surprised to find that a typical stock listed on a public stock exchange has a buy-and-hold return close to zero over its entire lifetime. Hence, the majority of stocks, which are risky assets, do not generate returns greater than the risk-free asset. This is contradictory to the financial theory that risky assets should earn a risk premium over the risk-free asset. However, Mehra and Prescott (1985) show that aggregating stock returns in a market return produces a positive and higher return than the return of the risk-free asset. Stock market returns are positive and higher than the return of the risk-free asset due to a small proportion of stocks that produce exceptional returns. Exactly how exceptional was recently shown by Bessembinder (2018), who calculated that just 4% of firms listed in the US was responsible for all net wealth creation in the US stock market from 1926 until 2016. On a global scale the asymmetry is larger. In a follow-up study covering 42 countries, Bessembinder et al. (2019) showed that just 1.3% of firms was responsible for all net wealth creation in the global stock market from 1990 until 2018[1]. Such severe asymmetry in long-term stock returns is detrimental for investors holding concentrated portfolios of stocks, as they might not be including those scarce long-term winning stocks. In fact, asymmetry in returns increases the probability of investors holding portfolios other than the market portfolio to experience lower returns than the market.

Asymmetry in long-term stock returns can be described with a more concise term, namely *positive skewness*. This means that the right tail in the distribution of long term stock returns is very long and thin, i.e. a few stocks have extremely positive returns and the majority have below average returns. Farago and Hjalmarsson (2019) simulate and test two strategies to reduce the drag of positive skewness on portfolio returns. First, portfolio diversification lowers the probability of missing stocks with exceptional long-term returns. Second, frequent rebalancing, i.e. changing the weights of portfolio stocks, opens the opportunity to achieve exceptional returns from stocks that themselves are not extreme winners. However, both diversification and rebalancing take skewness as a given and rely on adapting while it occurs. Both institutional and retail investors could, in theory, benefit from knowing what causes skewness and concentration in wealth creation. This allows them to anticipate and potentially avoid markets with extreme wealth concentration, thereby reducing the return drag skewness and wealth concentration have on their portfolios. Essentially, preventing instead of curing the disease.

However logical this may sound, literature about the causes of skewness and wealth concentration is scarce. Existing literature about stock returns predominantly focuses on monthly or annual returns. Yet, long-term returns are more likely to match the actual experience and needs of an investor, such as a pension fund having liabilities spanning decades. Some inroads have been made. Farago and Hjalmarsson (2019) show through simulation that skewness in long-term stock returns is induced by compounding and intensified by short-term volatility of returns. Additionally, Bessembinder et al. (2019) investigate differences in concentration of wealth creation

[1] Both papers define wealth creation of a stock as: "the difference between the present US dollar value of investors' actual investments in the stock and the value that would have been obtained if the same capital investments had earned one-month US Treasury Bill returns."

across countries and find a weak role for GDP related metrics but no role for other country characteristics such as individualism. Clear answers to what drives concentration in wealth creation and skewness have not been formulated. Additionally, research about skewness and concentration of wealth creation has been focused on the stock market as a whole. This paper differs from existing literature and contributes to it, by taking a novel approach of investigating skewness and concentration of wealth creation in industry segments of the US stock market. Hence, the research question of this paper is:

**Do US industries suffer from concentration in wealth creation and what are the empirical determinants of cross-industry variation in concentration of wealth creation?**

Thanks to segmentation of the stock market into 50 industries this paper is able to observe differences across industries in skewness, concentration in wealth creation and stock return performance against benchmarks such as the risk-free rate and inflation. This allows this paper to perform regression analysis by using concentration of wealth creation as the dependent variable and industry characteristics as the independent variables. This paper contributes by researching whether corporate performance and its variability across industries and within firms is a potential cause for stocks' underperformance of the risk-free rate and concentration in wealth creation. Including corporate performance establishes a link between the behaviour of stock returns and what is actually happening inside industries and firms. The use of corporate performance measures, such as profitability, financial distress and earnings volatility to explain the cross-section of stock returns is broadly established in literature by authors such as Novy-Marx (2013), George and Hwang (2010) and **?**. Please note, this paper does not seek to explain or predict asset prices and does not employ CAPM or factor models. We aim to uncover and explain cross-industry differences in concentration of wealth creation.

This paper takes into account a complete and realistic picture of common stock behaviour in the long-term. Therefore, it uses all available stock return data from CRSP in the period 1926 to 2018. Delisted stocks are included to prevent survivorship bias, making up almost 80% of the sample. Accounting data to construct financial ratios that measure corporate performance on the industry and firm level is retrieved from Compustat which is available from 1970 until 2018. Subsequently, buy-and-hold stock returns are compounded over annual, decade and lifetime horizons, while taking into account industry switching, to observe differences across industries in skewness. Buy-and-hold stock returns do not always reflect the experience of investors. Therefore, the amount of wealth created for investors in aggregate by individual stocks and industries is created along the lines of Bessembinder (2018) which allows this paper to measure the degree of concentration in wealth creation.

To explain differences in the cross-section of concentration in wealth creation, this paper utilizes fractional outcome regressions which are ideal for models where the dependent variable is a proportion (i.e. x% of firms is responsible for y% of wealth). Outcomes of the models are nonlinear. Therefore, marginal effects are plotted to intuitively understand what the expected effect is of the independent variables. Furthermore, a discussion can be held about how to measure wealth creation by stocks. Does the current method, based on the risk-free rate, match how institutional and retail investors experience wealth created by their investments? To provide an alternative perspective, this paper proposes a new method that is based on inflation, i.e. "real returns". In addition to testing hypotheses using the ordinary method of wealth creation this paper increases robustness by testing hypotheses with the alternative method.

This paper finds that large differences exist across US industries in terms of skewness of both short- and long-term stock returns. Skewness generally increases substantially with longer horizons. The mean return over lifetime horizons is positive for all industries, whereas the median stock return is negative in 23 out of 50 industries in the period from 1926 to 2018. Additionally, large variation in the ability of stocks within industries to outperform benchmarks such as the risk-free rate, rate of inflation and equal and value-weighted market return are found. For example, in the precious metals, recreation and computer industry more than 70% of stocks are unable to achieve a higher return than the risk-free rate over the full lifetime of the stock. On the other hand, in a few industries the majority of stocks outperform the risk-free and inflation rate, such as in the utilities, tobacco and candy & soda industry. Furthermore, not a single industry exists where the majority of stocks outperform either the equal- or value-weighted market return over their lifetime.

Large differences are present in absolute wealth creation across industries. Some industries create none or even destroy wealth in the entire period from 1926 until 2018 such as the coal, fabricated products and precious metals industry. On the contrary, the pharmaceutical, computer software and petroleum and natural gas industry create close to $3 trillion in wealth each during the period from 1926 until 2018. Regarding concentration of wealth creation, wide differences occur among industries. In some industries less than 1% or even a single firm is responsible for all net wealth creation throughout the period from 1926 to 2018. Industries with the most concentration in wealth creation are the electronic equipment, computer software and computer industry. With respect to the empirical determinants of concentration in wealth creation we find that short-term return volatility in a given industry is associated with more concentration of wealth creation. Hence, long-term investors should be aware that stock selection in industries with high short-term volatility is more difficult, because one is likely to miss the few stocks that create all the wealth. Better corporate performance of industries, higher variability in corporate performance within firms and higher short-term returns are generally associated with lower concentration in wealth creation. Therefore, long-term investors must be cautious if they attempt stock selection in industries with relatively bad corporate performance and low short-term mean returns. Variability in corporate performance of a given industry is often correlated with a low proportion of firms being able to outperform the risk-free rate. Hence, investors should be cautious in this environment too.

The remainder of this paper is structured as follows. Section 2 reviews existing literature on the topic of this paper. Section 3 develops the hypotheses to be answered. Section 4 provides summary statistics of the data and shows how the sample is constructed. Section 5 discusses the methodology and subsequent models that are used to answer the hypotheses. Section 6 reports empirical evidence and results for performance against benchmarks and regression analysis wealth concentration. Section 7 challenges the findings by proposing a new method for calculating wealth creation and checks if the findings hold. Section 8 discusses the limitations of this research and highlight opportunities for further research. Finally, section 9 sums up the findings and concludes the paper.

## 2 Literature review

Equity investors have a very large set of possible investments. They can choose to buy and sell thousands of stocks representing all sort of firms at any given time. To get a quick and easy understanding of all those stocks they can summarize them by taking averages. In other words, "stocks from this industry or with these characteristics on average have a x% return". Another option is to summarize the whole stock market by saying the market has a x% return annually. However, a dangerous mistake lurks in the fact that an average return obscures the underlying return distribution of a group of stocks. For example, a few stocks might have an exceptionally high return whereas most stocks have a mediocre return. In such a case, the average stock return would look promising, but in reality the majority of stocks have a rather disappointing return.

The widely received work of Bessembinder (2018) shows that long-term US stock returns are extremely positively in the period from 1926 until 2016. His work also shows that a very small proportion of stocks has a very high return over their lifetime and most other stocks have a negative or close to zero return over their lifetime. Hence, the median stock return is significantly lower than the mean stock return. Additionally, Bessembinder (2018) calculates wealth creation by taking the dollar return above the risk-free rate at present value for the full lifetime of each firm in the sample. He finds that just 4% of firms listed on the US stock market from 1926 until 2016 is responsible for 100% of wealth creation in the stock market.

Such a degree of concentration implies that the distribution of long-term stock returns is not normal, but is instead severely positively skewed. Simkowitz and Beedles (1978) investigate monthly holding period returns of all stocks listed on the New York Stock Exchange from 1945 until 1965 and empirically showed that individual stocks return are indeed positively skewed. Simulation evidence for stocks' return skewness is provided by Fama and French (2018), who show that investment payoffs deviate from returns that could be expected if returns were normally distributed. Farago and Hjalmarsson (2019) produce a theoretical framework and mathematically show that long-term stock returns are inherently skewed. Even more convincing evidence is provided by Bessembinder et al. (2019) who show that global stocks return distributions are even more skewed than US stock returns by doing the same exercise as in Bessembinder (2018) across 42 countries. They calculate that 1.3% of firms are responsible for 100% of the net wealth creation in the global stock market. In addition, they show that the majority of firms does not achieve a higher return than the risk-free asset, in this case US Treasury bills. Therefore, stock return skewness seems to complicate the work of investors, who are seeking those investments with the best risk-return relations on the long-term. Namely, excellent or even good long-term investments seem to be more rare than one might have considered before.

Without some stock picking skills an equity investor is likely to obtain a lower return than the market return unless he holds the whole market. Ikenberry et al. (1998) show through regression analysis that active fund managers, meaning those investment managers whose goal is to achieve a higher return than the market, experience a drag on returns due to skewness in individual stock returns. In a short simulation study, Heaton et al. (2017) show why indexing, a method of investing that mimics the return of the market, has a higher probability of success than selectively buying stocks. They find that asymmetry in stock return distributions causes stock pickers to frequently miss out on the relatively few stocks that perform exceptionally well. French (2008) shows how fees and trading costs significantly lower returns for an investor who chooses an "active" strategy, i.e. a strategy aiming to outperform the market return. In addition to those costs, skewness in stock returns is likely to give stock pickers and active managers an inherent disadvantage, because they are likely to miss long-term winners.

How can an equity investor reduce or mitigate the inherent disadvantage caused by skewness in stock returns? One way is to give up on stock picking and invest in passive products, i.e. products that mimic market returns such as index funds. Passive investing has gained popularity in recent years. However, Blitz (2014) points out that adoption of passive investing by the whole market is undesirable, because some investors still need to guide the way for passive investors by actively allocating capital. Additionally, research by Kraus and Litzenberger (1976) shows that investor actually prefer positive skewness. Therefore, one could state that investors are naturally drawn to stock picking, since Neuberger and Payne (2018) show that market returns (which one would obtain by indexing) are negatively skewed. Hence, simply switching to passive investing is not going to do the trick for everyone.

Consequently, an investor still wishing to actively select stocks must find strategies to reduce the disadvantage caused by stock return skewness. One such strategy is diversification. Simkowitz and Beedles (1978) construct portfolios with increasingly more randomly selected stocks and calculates their attributes, essentially diversifying stock portfolios. They show that portfolios containing more stocks suffer less from individual stock return skewness. This strategy is confirmed by Farago and Hjalmarsson (2019), whom use their theoretical framework and simulation to show the power of diversification in mitigating skewness. Additionally, they give an important role to rebalancing stock portfolios, i.e. increasing or decreasing the weight of stock in the portfolio. They show through simulation that even relative infrequent rebalancing, every 5 years over a 30 year investment horizon, improves the probability of outperforming the market return with a concentrated portfolio. Diversification and rebalancing are two relatively simple strategies reducing the inherent disadvantage of an active investor.

Instead of taking the disadvantage connected to positive skewness as a given, investors could potentially benefit knowing the causes of skewness and the associated concentration of wealth creation. How is it possible that only 4% of companies is responsible for 100% of the wealth creation in the stock market? Arditti and Levy (1975) were one of the first to recognize the effects of compounding when he disagreed with earlier authors such as Friend and Blume (1970). Friend and Blume (1970) argued that the skewness of a stock portfolio should be zero when monthly returns are used. Bessembinder (2018) adds to the arguments put forward by Arditti and Levy (1975). In his work he draws monthly returns from normal distributions with different standard deviations. Subsequently, by linking monthly returns he creates buy-and-hold returns for annual, decade and a 90-year horizon and finds that skewness increases with higher levels of return standard deviation. In the same simulation exercise Bessembinder (2018) concludes that compounding of single-period returns induces skewness over longer horizons and causing the median buy-and-hold return to be lower than the mean return. Farago and Hjalmarsson (2019) perform a similar exercise and also conclude that compounding of returns drives skewness. They also find that short-term standard deviation, i.e. short-term volatility of returns increases skewness in long-term stock returns.

To grasp extreme skewness and concentration of wealth creation one could investigate whether these phenomena occur also occur in selected groupings of stocks, instead of the market as a whole. Bessembinder et al. (2019) were the first and only ones until now to delve into this question. The authors collected a very large sample of stock returns from 42 countries from 1990 to 2018 and observed that the degree of wealth concentration and the level of skewness varies considerably across countries. For example, the top 0.5% firms in Spain created 30% of net wealth in the Spanish stock market, whereas the top 0.5% firms in Indonesia created 80% of net wealth. They also observed that the proportion of stocks that is able to outperform the risk-free asset (US Treasury bills) varies significantly across countries. In an attempt to explain these differences they performed regression analysis using technical (i.e. volatility) and country (i.e. GDP, individualism) characteristics as independent variables. The authors found convincing evidence for short-term volatility and weaker evidence for GDP/capital and GDP growth to be associated with concentration in wealth creation. One of the problems might have been the lack of variation in country specific variables or the weak relation with stock market returns. Therefore, the puzzle of stock return skewness, the inability of most stocks to outperform the risk-free rate and concentration of wealth creation is not solved yet and more research is necessary.

# 3 Hypothesis development

An attractive research gap emerges from the puzzle of stock return skewness: the inability of most stocks to outperform the risk-free rate in combination with concentration of wealth creation. In an attempt to solve that puzzle this paper takes a novel approach by comparing industries within the United States. Grouping stocks according to their industries allows this research to obtain a more detailed view of where stock return skewness, outperformance of the risk-free rate and concentration of wealth creation occur. On top of this, grouping stocks into industries enables us to exploit differences between industries to help explain what drives stock return skewness, outperformance of the risk-free rate and concentration of wealth creation.

The first hypothesis aims to answer the basic question whether positive skewness of stock returns is present across US industries and whether there are differences in the level of skewness across US industries. Positive skewness means that a few stocks have exceptional returns and most stocks have below-average returns. If positive skewness is present it is likely that wealth creation is concentrated since only a few stocks have very high returns. Hence, investigating skewness in stock returns helps to answer the first part of the research question: "Do US industries suffer from concentration in wealth creation?" Based on research by Bessembinder (2018) and Farago and Hjalmarsson (2019) we expect stocks returns in US industries to be positively skewed. The authors of both papers show through simulation and historical evidence that on a country level long-term stock returns are skewed. Hence, the first hypothesis is:

**Hypothesis 1:** long-term buy-and-hold stock returns across US industries are positively skewed and differences in skewness exist across industries.

The second hypothesis also aims to answer the first part of the research question: "do US industries suffer from concentration in wealth creation?" Remember that wealth is created as long as stocks achieve a higher return than the risk-free asset. Hence, if a large proportion of stocks is unable to outperform the risk-free rate it means that only a small proportion of stocks create wealth. If only a small proportion of stocks create wealth, wealth creation would be concentrated by definition. This paper suspects that the majority of stocks across US industries is unable to outperform the risk-free rate. Both Bessembinder (2018) and Bessembinder et al. (2019) analyze individual stock returns on a country level and find that the majority of stocks is unable to outperform the risk-free rate over their lifetime. This paper contributes by assessing whether stocks from a given industry are able to outperform the risk-free rate. Hence, the second hypothesis is:

**Hypothesis 2:** the majority ($> 50\%$) of stocks across US industries is unable to achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset.

By measuring what proportion of stocks in a given industry is able to outperform the risk-free rate we are likely to find that differences in that proportion exist across industries. For example, Bessembinder et al. (2019) find that large differences exist in the proportion of stocks that is able to outperform the risk-free rate across countries. Although we measure whether an individual stock is able to outperform the risk-free rate, industry level returns from Kenneth French's data library[2] already show that large differences occur in returns and thus the proportion of stocks that outperform the risk-free rate could also differ across industries. If these differences in the ability to outperform the risk-free rate are present it is an indicator that concentration in wealth creation also varies across industries. Hence, researching cross-industry variation in the proportion of stocks that is able to outperform the risk-free rate helps to answer the second part of the research question: "what are the empirical determinants of cross-industry variation in concentration of wealth creation?". Hypotheses 3, 4 and 5 present three sets of empirical determinants that might have the ability to explain cross-industry variation in the proportion of stocks that outperform the risk-free rate.

**Hypothesis 3:** short-term stock return volatility, stock return skewness and mean stock returns can explain differences in the proportion of stocks across US industries that achieve a higher

[2] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

buy-and-hold return over their lifetime than the return of the risk-free asset.

First, research by Bessembinder et al. (2019) found that short-term stock return volatility is a significant predictor for explaining the proportion of stocks in each country that outperform the risk-free rate. Specifically, Bessembinder et al. (2019) found that short-term stock return volatility was negatively associated with the proportion of stocks that outperformed US Treasury bills over their lifetime. Bessembinder (2018) and Farago and Hjalmarsson (2019) provide an explanation for this result by assuming short-term stock returns are distributed lognormally. They mathematically show that skewness in long-term stock returns increases dramatically and non-linearly as standard deviations of short-term stock returns are higher. Skewness in long-term stock returns implies that a large proportion of stocks have below-average returns. In extreme cases of skewness the median return can be much lower than the mean return. For example, an extremely positively skewed stock return distribution could have a median return of 0% and a mean of 6% as Bessembinder (2018) shows in his simulation exercise. Hence, skewness could be so extreme that a large proportion of stocks have a lower return than the risk-free rate. Thus, this paper expects that short-term stock return volatility, which induces skewness, is negatively associated with the proportion of stocks that outperform the risk-free rate over their lifetime in a given industry.

Second, this paper also expects that short-term stock return skewness can explain cross-industry variation in the proportion of stocks that achieve a higher buy-and-hold return over their lifetime than the risk-free rate. Conrad et al. (2013) test if stocks' short-term skewness is related to future returns and find that a higher short-term skewness is strongly related to lower future returns. Additionally, Farago and Hjalmarsson (2019) find that short-term skewness of stock returns increases long-term skewness if short-term volatility is high. Thus, this paper expects that short-term return skewness is negatively associated with the proportion of stocks that outperform the risk-free rate over their lifetime in a given industry.

Last, this paper expects that a higher short-term mean stock return in a given industry is associated with a higher proportion of stocks that outperform the risk-free rate over their lifetime. It is intuitive that if short-term returns are higher long-term returns are likely to be higher. Thus, the probability of a stock outperforming the risk-free rate over its lifetime could be higher if short-term stock returns are higher. In conclusion, hypothesis 3 conjectures that short-term return volatility, stock return skewness and mean stock returns are empirical determinants of cross-industry variation in the proportion of stocks that outperform the risk-free rate.

The empirical determinants in hypothesis 3 focus on technical and short-term characteristics of stock returns. However, determinants with a more direct economic link to firms inside industries might also be able to explain cross-industry variation in the proportion of stocks that exceed the return of the risk-free asset over their lifetime. This paper aims to establish such a direct economic link by using corporate performance of firms within the industries we research. For example, an industry where more firms have higher profitability (i.e. a higher median profitability) could theoretically have more firms that create wealth. In other words, profitability could be positively associated with the proportion of stocks that outperform the risk-free rate over their lifetime. In practice, a positive relation between profitability and stock returns was found by Fama and French (2015) and Novy-Marx (2013). In this example, profitability is hypothesized to be an empirical determinant of cross-industry variation in the proportion of stocks that outperform the risk-free rate over their lifetime. We aim to research whether corporate performance plays a role in concentration of wealth creation. Hence, hypothesis 4 of this paper is:

**Hypothesis 4:** corporate performance can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset.

It is important to understand that this paper does not seek to explain cross-sectional differences in stock returns by using "factors" such as size, value and momentum which are researched in much cited work by Fama and French (1992) and Jegadeesh and Titman (1993). Instead, this paper focuses on concentration of wealth creation and looks at industries and the market as a whole over very long time horizons. We do not formulate a trading strategy by taking long positions in stocks with favourable characteristics and short positions in stocks with unfavourable characteristics as is done in factor investing. However, we do acknowledge that findings from the class of factor investing research can be helpful in developing a hypothesis on how corporate performance could influence concentration of wealth creation.

This paper expects better median corporate performance of a given industry to be associated with a larger proportion of stocks that outperform the risk-free asset over their lifetime. Five categories of corporate performance are used to measure the "business performance" of US firms within their industries. These five categories are capitalization, solvency, liquidity, efficiency and profitability.

First, capitalization measures how much debt a firm uses relative to its total capital structure. Ohlson (1980) researches how financial ratios influence a firm's probability of default and finds that a higher relative amount of debt is associated with a higher probability of default. Industries containing firms with a higher probability of default could be prone to wealth destruction and thus have a lower proportion of firms outperform the risk-free rate over their lifetime. In addition, George and Hwang (2010) assess US stock returns from 1965 to 2003 and find that the relation between stock returns and leverage (i.e. relatively high debt) is significantly negative. Therefore, this paper expects that a higher capitalization ratio, i.e. the use of more debt, in a given industry is associated with a lower proportion of stocks that outperform the risk-free rate.

Second, solvency captures a firm's ability to meet long-term obligations. Beaver (1966) finds that cash flow to debt ratios have the ability to predict failure of firms by researching US firms from a wide range of industries. Therefore, this paper expects that better solvency in a given industry is associated with a higher proportion of stocks that is able to outperform the risk-free rate over their lifetime.

Third, liquidity captures a firm's ability to meet short-term obligations such as payments to suppliers or repayment of short-term debt. Low liquidity could indicate that a firm finds itself in financial distress, because it might be unable to honor short-term contracts. Piotroski (2000) investigates value stocks in the United States and finds that portfolio returns increase significantly if financially distressed firm are excluded. Altman (1968) constructs an accurate scoring system to predict corporate failure in manufacturing firms and finds that liquidity is an important component in determining the probability of default. Therefore, this paper expects that higher liquidity in a given industry is associated with a higher proportion of stocks that is able to outperform the risk-free rate over their lifetime.

Fourth, efficiency measures how effective firms use assets and liabilities. For example, a firm that is able to generate a large amount of sales compared to the amount of assets (asset turnover) it uses is considered to be efficient. Soliman (2008) researches how market participants use DuPont analysis, of which asset turnover is a component, and finds that annual stock returns are positively influenced by an increase in asset turnover. Haugen et al. (1996) conjecture that higher asset turnovers indicate larger growth potential of firms and should positively contribute to returns. However, Hou et al. (2017) test asset turnover as a determinant for future returns in the US and show it is insignificant. Although evidence is mixed this paper expects that more efficient industries have a larger proportion of stocks that are able to outperform the risk-free rate over their lifetime.

Last, profitability captures a firm's ability to generate profit. For example, higher operating margins imply that firms are able to generate more profit from their operations. Numerous studies have highlighted that higher profitability is associated with higher excess stock returns, i.e. returns above the risk-free rate. For example, Novy-Marx (2013) researches US stocks from 1963 to 2010 and finds that profitable firms generate significantly higher returns than unprofitable firms. In addition, Fama and French (2015) added profitability and investment to their well-known three-factor model (Fama and French (1992)) which explains cross-sectional returns in the US stock market. Fama and French (2015) found that high operating profitability is associated with higher average stock returns. Therefore, this paper expects that higher profitability in a given industry is associated with a larger proportion of stocks that outperform the risk-free rate over their lifetime. In conclusion, we expect that better corporate performance in terms of capitalization, solvency, liquidity, efficiency and profitability in a given industry is associated with a larger proportion of firms achieving a higher buy-and-hold return over their lifetime than the return of the risk-free asset.

This research takes a cross-sectional approach which means that we measure concentration of wealth creation and the ability to outperform the risk-free rate of industries once, as we are interested in long-term behaviour of stock returns. However, measuring median corporate performance of industries over the full length of the sample might conceal that corporate performance within firms of a given industry varies throughout time. We hypothesize that variability in corporate performance could contain important information about the degree of competition and the stability of firms in industries and thus their ability to create wealth. Dambolena and Khoury (1980) investigate how stability in corporate performance measures such as profitability, leverage and liquidity can predict corporate failure, i.e. bankruptcy. They find that stability of financial ratios is negatively associated with corporate failure. Betts and Belhoul (1987) investigate firm failure from 1974 to 1978 in the United Kingdom and find that including stability of financial ratios improved the ability to identify failed and non-failed firm. We expect that variability of corporate performance in a given industry is negatively associated with the proportion of stocks that exceed the return of the risk-free asset over their lifetime. Hence, hypothesis 5 of this paper is:

**Hypothesis 5:** variability in corporate performance can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset.

Hypotheses 2, 3, 4 and 5 indirectly answer the research question by investigating cross-industry variation in the proportion of stocks that is able to outperform the risk-free rate over their lifetime. Hypotheses 6, 7, 8 and 9 investigate concentration in wealth creation directly. Hypothesis 6 covers the first part of the research question: "do US industries suffer from concentration in wealth creation?". Hence, hypothesis 6 of this paper is:

**Hypothesis 6:** a high degree of concentration in wealth creation, i.e. $< 5\%$ of stocks are responsible for 100% of wealth creation, is present across US industries.

Existing literature has not specifically investigated concentration of wealth creation in US industries. However, Bessembinder (2018) and Bessembinder et al. (2019) report which firms were responsible for large portions of total wealth creation. Based on visual inspection of their results, we expect that most industries suffer from concentration of wealth creation as just a few firms from every industry appear to create substantial wealth. By answering hypothesis 6 we will come to know whether that expectation is correct. We also expect that there are differences in concentration of wealth creation, because Grullon et al. (2019) shows that intensity of competition differs across US industries. Competition might limit the ability of a few firms to dominate an industry.

Hypotheses 7, 8 and 9 cover the second part of the research question directly: "what are the empirical determinants of cross-industry variation in concentration of wealth creation?". Hypotheses 7, 8 and 9 use the same determinants as hypotheses 3, 4 and 5. Logically, we expect opposite signs for hypotheses 7, 8 and 9 compared to hypotheses 3, 4 and 5, because if more firms outperform the risk-free rate over their lifetime we expect a lower concentration in wealth creation. Hypothesis 7 of this paper is:

**Hypothesis 7:** short-term stock return volatility, return skewness and mean returns can explain differences in the proportion of stocks that is responsible for a $x\%$ of wealth across US industries.

Why we expect short-term stock return volatility, return skewness and mean returns to have explanatory power for concentration of wealth creation is elaborately discussed for hypothesis 3. We briefly summarize the explanations from hypothesis 3. Short-term return volatility induces positive skewness of long-term returns which means that a few firms have exceptional returns (i.e. create substantial amounts of wealth) whereas most firms have below-average returns. Thus, we expect that short-term return volatility increases concentration in long-term wealth creation. Short-term return skewness could amplify long-term return skewness thereby increasing concentration of wealth creation. A higher short-term mean return could imply that long-term returns are also higher and thus decrease concentration of wealth creation.

**Hypothesis 8:** corporate performance can explain differences in the proportion of stocks that is responsible for a $x\%$ of wealth across US industries.

The discussion of hypothesis 4 provides a detailed explanation of why one can expect that better corporate performance is associated with more firms outperforming the risk-free rate over their lifetime. Hence, we expect that better corporate performance in terms of capitalization, solvency, liquidity, efficiency and profitability in a given industry is associated with a lower degree of concentration in wealth creation.

**Hypothesis 9:** variability in corporate performance can explain differences in the proportion of stocks that is responsible for a $x\%$ of wealth across US industries.

The development of hypothesis 5 explains that variability in corporate performance of firms in a given industry could indicate that firms are more prone to go bankrupt and thus destroy wealth. In addition, Bessembinder (2018) find that underporformance of the risk-free rate by the majority of stocks is primarily driven by stocks that are delisted by stock exchanges. However, variability in corporate performance could in theory indicate that an industry is highly competitive although specific research supporting that theory has not been published yet. Hence, the expected relation between variability of corporate performance and wealth concentration cannot be inferred from existing literature. However, common sense of the author would say that industries where firms experience high variability in corporate performance contain firms that are unstable or suffer from intense competition. Therefore, we expect that variability in corporate performance is negatively associated with concentration of wealth creation.

# 4 Data

## 4.1 Stock returns

Stock returns are a key component to answer the hypotheses of this research. Long-term stock returns are calculated to test hypotheses about skewness and stock return performance against return of the risk-free asset. Hypothesis 1 states that long-term buy-and-hold stock returns across US industries are positively skewed and differences in skewness exist across industries. Hypotheses 2, 3, 4 and 5 make claims about stocks' ability to achieve a higher return than the risk-free rate over their lifetime. Hypotheses 6, 7, 8 and 9 make claims about concentration of wealth creation. Monthly stock returns are a component in wealth creation calculations. Therefore, reliable stock return data is paramount in answering all hypotheses. This section describes which stock returns are used, where they were retrieved and how monthly stock returns are compounded to create long-term stock returns.

This paper chooses to investigate the hypotheses in the United States, because data availability is much better in terms of historical returns and accounting figures. In order to construct return series for individual stocks, Center for Research in Security Prices (CRSP) data is used. This paper seeks to capture the US stock market in its entirety and all its extremities. Therefore, all common stock (share code 10, 11 and 12) and real estate investment trust (share code 18) data from January 1926 until December 2018 is extracted from the CRSP database. Real estate investment trusts (REITs) are included, because they form a substantial part of the equity market[3] and because their classification as real estate or equities is debatable (Morawski et al. (2008)). We cover the full length of the CRSP database stretching over 1116 months or 93 years. The frequency of the data is monthly. The total number of unique stocks in the sample is 26,993.

Return data incorporates price, dividends and delisting returns and is equal to the monthly buy-and-hold return of an investor. Long-term stock returns are calculated by compounding monthly returns, $R_t$, over the desired period from $t$ until $T$. For example, $R_{t,T}$ can be calculated using the formula:

$$R_{t,T} = (1 + R_{t+1}) * (1 + R_{t+2}) * ... * (1 + R_T) - 1 \tag{1}$$

However, for large samples such as the one this paper uses, summation of logarithmic returns is more efficient than multiplication of simple returns. Furthermore, this paper uses the statistical software package Stata which is programmed in a way that is does not handle the multiplication method described above well. Therefore, this paper uses two properties of natural logarithms to compound monthly returns.

By making use of natural logarithms one can use summation instead of multiplication:

$$ln(x * y) = ln(x) + ln(y) \tag{2}$$

Using the return series from (1) this translates to:

$$ln((1 + R_{t+1}) * ... * (1 + R_T)) = ln(1 + R_{t+1}) + ... + ln(1 + R_T) \tag{3}$$

Which can be rewritten to:

$$ln(1 + R_{t,T}) = ln(1 + R_{t+1}) + ... + ln(1 + R_T) \tag{4}$$

Euler's number, $e$, enables one to go back from the natural logarithm environment:

$$e^{ln(x)} = x \tag{5}$$

Which for a return series translates to:

$$e^{ln(1+R_{t,T})} = 1 + R_{t,T} \tag{6}$$

[3] https://www.reit.com/data-research/reit-market-data/us-reit-industry-equity-market-cap

Hence, to calculate returns over longer time horizons, this paper takes the sum of monthly returns, $ln(1+R_t)$, over the desired time horizon and translates it back by using the exponential function $e^{ln(x)}$. The full formula is depicted as:

$$R_{t,T} = e^{ln(1+R_{t+1})+...+ln(1+R_T)} - 1 \tag{7}$$

The method described above is used to calculate annual, decade and lifetime buy-and-hold returns from monthly buy-and-hold returns. These returns are used to investigate whether buy-and-hold returns across US industries are indeed positively skewed as presumed by hypothesis 1. To calculate returns over annual horizons, calendar years are used to prevent overlap. Decade returns are calculated using fixed decades with the first decade running from calendar years 1926-1938 (13 years) and all consecutive decades running for 10 calendar years ending in 2018. Lifetime returns are calculated from the first appearance in the sample until the last appearance. Lifetime returns are used to investigate hypotheses 2, 3, 4 and 5, which make claims about stock's ability to achieve a higher return than the risk-free asset during their lifetime.

## 4.2 Assigning stocks to industries

All hypotheses have an element of comparing industries with each other or use differences across industries to explain variation in performance against the risk-free asset and wealth concentration. Therefore, this paper categorizes stock observations into industries on a monthly basis. Individual stock observations are assigned to an industry based on their Standard Industry Classification (SIC) code. A two stage approach is taken.

First, CRSP historical SIC codes are used as the main source of SIC codes, which covers 96.32% or 3.67 million out of 3.81 million observations in the sample. If CRSP historical SIC codes are not available, the most recent SIC code from CRSP is used, this covers an additional 90 thousand observations. In the case that a CRSP SIC code is unavailable, Compustat SIC codes are used. This covers an additional 53 thousand observations. The last option is to use a SIC crosswalk from NAICS codes, which covers 439 observations. In total, 99.92% of the observations in the sample are assigned a SIC code.

Second, this paper uses the industry classification system Kenneth French[4] designed, which is based on SIC codes. The system assigns ranges of SIC codes to certain industries. The least detailed system assigns all firms to a total of five industries, whereas the most detailed system assigns firms to 49 industries. This paper chooses to use the most detailed system of 49 industries to diminish the chance that firms with unrelated product markets or business models are allocated to the same industry. This paper thankfully makes use of the Stata package coded by Alfen (2017) that automatically assigns observations with a certain SIC code to an industry based on Kenneth French's industry classification systems.

Although practically every observation has an SIC code, some SIC codes cannot be assigned to an industry. This is the case for 19,987 observations or 0.52% of the sample. These observations are

[4] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

classified under the industry group ”Not available” which brings the total number of industries to 50. Table 1 displays the industry summary statistics for the full sample on a monthly basis. Because accounting data used in hypotheses 3 until 5 and 7 until 9 is only available in the period 1970 to 2018 we also display separate summary statistics for this period in Table 1.

Table 1: Frequency table of monthly observations per industry for 1926-2018 and 1970-2018

| | Observations (1926-2018) | Observations (1970-2018) |
|---|---|---|
| **Fama-French 49 Industries** | | |
| Agriculture | 12,639 | 11,325 |
| Food Products | 82,644 | 49,983 |
| Candy & Soda | 10,898 | 9,435 |
| Beer & Liquor | 17,482 | 11,773 |
| Tobacco Products | 13,538 | 4,645 |
| Recreation | 33,896 | 29,229 |
| Entertainment | 48,003 | 41,040 |
| Printing & Publishing | 36,169 | 32,377 |
| Consumer Goods | 78,930 | 62,715 |
| Apparel | 54,324 | 41,094 |
| Healthcare | 54,267 | 54,235 |
| Medical Equipment | 82,708 | 80,300 |
| Pharmaceutical Products | 134,103 | 126,592 |
| Chemicals | 76,170 | 54,015 |
| Rubber and Plastic Products | 29,169 | 25,845 |
| Textiles | 39,744 | 25,241 |
| Construction Materials | 112,071 | 80,836 |
| Construction | 40,165 | 36,598 |
| Steel Works | 82,203 | 44,138 |
| Fabricated Products | 13,446 | 12,791 |
| Machinery | 132,358 | 100,449 |
| Electrical Equipment | 87,520 | 72,053 |
| Automobiles and Trucks | 71,673 | 43,569 |
| Aircraft | 25,407 | 16,152 |
| Shipbuilding, Railroad Equipment | 12,618 | 5,473 |
| Defense | 5,606 | 5,248 |
| Precious Metals | 36,682 | 35,024 |
| Mines | 37,180 | 25,599 |
| Coal | 11,288 | 5,715 |
| Petroleum and Natural Gas | 180,573 | 151,850 |
| Utilities | 128,811 | 95,553 |
| Communication | 81,069 | 75,464 |
| Personal Services | 34,575 | 32,499 |
| Business Services | 185,549 | 178,404 |
| Computers | 88,591 | 82,629 |
| Computer Software | 143,149 | 143,090 |
| Electronic Equipment | 161,641 | 151,875 |
| Measuring and Control Equipment | 63,753 | 60,370 |
| Business Supplies | 35,003 | 30,702 |
| Shipping Containers | 27,565 | 15,973 |
| Transportation | 116,806 | 74,431 |
| Wholesale | 129,035 | 120,724 |
| Retail | 193,126 | 154,066 |
| Restaurant, Hotels, Motels | 70,724 | 65,252 |
| Banking | 249,167 | 241,656 |
| Insurance | 94,533 | 92,184 |
| Real Estate | 94,741 | 90,598 |
| Trading | 233,353 | 221,326 |
| Other | 12,586 | 11,974 |
| Not available | 19,987 | 17,667 |
| Total | 3,817,238 | 3,251,776 |

The first sample ranges from January 1926 until December 2018. The second sample ranges from January 1970 until December 2018. Source: CRSP, Compustat and Kenneth French data library.

### 4.2.1 Stocks switching industries

Since this paper concerns itself with long-term returns across industries a system is needed that allows stocks to switch between industries. For example, a firm that used to make computer hardware could switch to designing computer software. In that case the stock switches from Kenneth French's industry code 35 (computer hardware) to 36 (computer software). Some stocks switch to another industry once and stay there. However, other stocks switch back and forth between industries. Within the sample 7,975 industry switches occur on a total number of 26,993 unique stocks.

The full sample used in this paper is a panel data set, with a unique stock identifier from CRSP, PERMNO, serving as the panel identifier and a monthly date serving as the time identifier. Within these panels every monthly observation is assigned to an industry as described in the previous section. To account for industry switching, this paper creates another panel within the stock panels using the industry identifier as the panel identifier. Table 2 serves as an example. It is important to note that a new panel is created when the stock switches back to an industry it was previously in. Otherwise, non-continuous monthly returns are linked resulting in miscalculated long-term returns.

Table 2: Panel within panel data example

| PERMNO | Date | Return | Industry | Panel ID |
|---|---|---|---|---|
| 1 | January 2018 | 2% | Banking | 1 |
| 1 | February 2018 | 5% | Banking | 1 |
| 1 | March 2018 | -3% | Insurance | 2 |
| 1 | April 2018 | 4% | Insurance | 2 |
| 1 | May 2018 | -6% | Banking | 3 |
| 1 | June 2018 | 3% | Banking | 3 |

In a more complex scenario, where one is interested in yearly returns across industries a yearly panel must be created within the panel dataset. See Table 3 as an example. A panel ID is generated based on the calendar year and the industry. The same method is used for decade and lifetime returns.

Table 3: Yearly panel within panel data example

| PERMNO | Date | Year | Return | Industry | Panel ID |
|---|---|---|---|---|---|
| 1 | January 2005 | 2005 | 2% | Banking | 1 |
| 1 | ... | 2005 | ... | Banking | 1 |
| 1 | January 2006 | 2006 | -3% | Banking | 2 |
| 1 | ... | 2006 | ... | Banking | 2 |
| 1 | August 2006 | 2006 | 5% | Insurance | 3 |
| 1 | ... | 2006 | ... | Insurance | 3 |
| 1 | January 2007 | 2007 | 3% | Insurance | 4 |
| 1 | ... | 2007 | ... | Insurance | 4 |

As can be seen in Table 3, a panel can span over a shorter time horizon than the horizon one is analyzing. Panel ID 2 and 3 span 6 months each. When calculating return statistics across industries this could bias the results if a large number of panels are shorter than the intended time horizon. This paper investigates if this occurs and finds that there is no significant bias in annual returns. Regarding decade returns, a full decade lasts 120 months. However, the mean number of months for decade panels is 60 months in the sample and variation in the number of months is present across industries. In section 4.3, statistics on the length of panels across industries for decade and lifetime returns are reported to inform the reader about potential bias.

## 4.3 Summary statistics stock returns

This section describes summary statistics for monthly, annual, decade and lifetime buy-and-hold stock returns. Stocks are grouped in industries and industry switching is taken into account as described in the previous section. Although this is not the empirical results part yet, we can test hypothesis 1 already. The hypothesis states that long-term buy-and-hold stock returns across US industries are positively skewed and differences in skewness exist across US industries. Therefore, the following sections report and explain if and how much skewness is found in buy-and-hold stock return on a monthly, annual, decade and lifetime horizon.

### 4.3.1 Monthly horizon

Skewness in stocks' buy-and-hold returns is already detected at relatively short horizons as shown in the sample with the period 1926 until 2018 which is summarized in Table 4. On a monthly horizon skewness is 6.439, thus the right-tail of the distribution is significantly longer than the left tail. All industries are characterised by positive skewness in their distribution of monthly buy-and-hold stock returns. Although monthly horizons cannot be classified as long-term, the findings hint that we might fail to reject hypothesis 1 at a later stage. A few extreme positive returns increase the mean monthly buy-and-hold return to 1.07% well above the median return of 0%. The left tail has a minimum return of -100% due to limited liability of equities which contributes to skewness. When stocks reach a price of 0 or close to zero they are often delisted by the exchange and thus cease to exist for the remainder of our sample. However, during their lifetime they are included in all calculations regarding skewness, outperformance of the risk-free rate and concentration of wealth creation to prevent survivorship bias.

Differences in skewness across industries is found at relatively short monthly horizons. This is an indication that we might fail to reject hypothesis 1. The medical equipment (18.8) and steel works (22.1) industry stand out by experiencing very high skewness. Visual inspection of the data clarifies that a few very high monthly returns occur in very small stocks in these industries, which causes the observed skewness. Particularly low skewness is observed in the defense (1.5) industry. 42 out of 50 industries had a median return of 0% on a monthly basis. No industry had a negative monthly median return. Which leaves 8 industries with a positive median monthly return. Mean returns vary substantially, the aircraft industry and measuring & control equipment industry posted the highest mean monthly buy-and-hold return of 1.40% and 1.39% respectively from 1926 to 2018. The lowest mean monthly return was recorded by the "Not available" group, visual inspection of the sample indicates that this group contains a substantial amount of very small and quickly failing stocks. Other industries that recorded a rather low mean monthly return are the agriculture (0.71%) and recreation (0.72%) industry. Table 25 in Appendix A displays summary statistics for the period 1970 to 2018, which are largely similar to the period 1926 to 2018 due to the large overlap.

#### 4.3.2 Annual horizon

Table 5 shows that positive skewness becomes more pronounced over an annual horizon. The skewness of annual buy-and-hold individual stock returns is 20.211 in the period 1926 to 2018. The variation in individual industry skewness also increases over longer horizons. Two industries that stand out are the restaurant, hotel and motels (skewness: 52.890) and the trading (skewness: 76.170) industry. On the contrary, industries such as coal (skewness: 1.049) and fabricated products (skewness: 1.969) show more modest positive skewness. Buy-and-hold stock returns are increasingly positively skewed in most industries and differences in skewness across industries are increasing. However, stocks in some industries exhibit a slightly lower skewness over annual horizons compared to monthly horizons. Therefore, evidence in support of hypothesis 1 is strong but not definite yet.

Mean and median returns show significant variation between industries. An investor randomly choosing one stock in the pharmaceutical industry was most likely to end up with a return of -0.14%. Whereas if he or she had held the whole industry equally-weighted, the return would have been 14.81%. This already indicates that an investor must possess significant stock picking skills when choosing to hold a concentrated portfolio. The utilities industry is an example where the mean (13.32%) and median (11.17%) return are quite close. Mean returns vary across industries, a stock in the agriculture industry returns 7.65% annually on average, whereas a stock in the defense industry appreciated 21.27% on average. Table 26 in Appendix A displays summary statistics for the period 1970 to 2018, which are largely similar to the period 1926 to 2018 due to the large overlap.

#### 4.3.3 Decade horizon

The effects of compounding become very clear in buy-and-hold returns of individual stocks over the decade horizon which are reported in Table 6. Skewness for the full sample in the period 1926 to 2018 rises to 29.7. Skewness is positive across all industries and large differences in skewness exist when looking at decade horizons. For example, skewness is 28.7 in the pharmaceutical industry whereas skewness is 2.3 in the candy & soda industry. These findings support hypothesis 1 which states that long-term buy-and-hold stock returns across US industries are positively skewed and differences in skewness exist across industries. These findings also provide an indication for the first part of the research question: "do US industries suffer from concentration in wealth creation?" Skewness is substantially positive for most industries which indicates that a few firms experience exceptional return whereas the majority of stocks experience below-average returns. Hence, concentration of wealth creation in US industries is likely to be present.

A decade yielding particularly extreme returns was the decade leading up to the dotcom bubble from 1989 to 1998. To name a few examples, Dell Computer recorded a 35,030% buy-and hold return, America Online Inc. recorded a 32,856% buy-and-hold return and Cisco Systems recorded a 28,642% buy-and-hold return. In the subsequent decade from 1999 until 2008 the dotcom bubble burst, Dell Computer recorded a buy-and-hold return of -72.02%, Cisco Systems returned -75.09% and America Online Inc. fell 72.82%. The median decade buy-and-hold return for all individual stocks in the sample is 16.67% whereas the mean return is 122.57%. It should not come as a surprise that skewness is particularly high in industries linked to the dotcom bubble. Computer, computer software and business services all show very high skewness. It is evident that there are a few winners among many losers in these industries as the median returns are negative while the mean returns are positive and relatively high compared to other industries possible due to the eventual success of a few companies such as Apple and Microsoft.

Table 5 also reports the mean time in months firms of a given industry spend in the sample. Note that firms can occur multiple times if they switch industries. If firms in particular industries switch more often, the mean time in the industry will be shorter. Time in an industry can also be shorter due to a firm being listed or delisted during a decade. Furthermore, mean and median buy-and-hold return differences can be partly explained by the duration a firm is in a particular industry and/or listed on the stock exchange, because compounding over longer horizons can yield extreme returns. Two industries that are characterised by particularly high barriers to entry stand out with respect to stocks' mean time in the industry, namely the tobacco and utilities industry.

Table 26 in Appendix A displays summary statistics for the period 1970 to 2018. Decade buy-and-hold returns are generally not higher or lower compared to the period 1926 to 2018. Skewness is also similar.

#### 4.3.4 Lifetime horizon

As depicted in Table 7, skewness is highest for lifetime buy-and-hold returns at 108.5 in the period from 1926 to 2018. It is evident that we accept hypothesis 1. Table 7 shows that long-term buy-and-hold stock returns across US industries are indeed positively skewed and differences in skewness across industries do exist.

Quite dramatically, the median return of an individual stock over its lifetime was 2% for the full sample. This might sound dire, but in 11 out of 50 industries the median buy-and-hold stock return was lower than -25%. How often does an investment banker tell a prospective investor that the most common historical outcome of an investment in the computer software industry was -50%? Probably, most rational investors would take their money and run. Variation in skewness across industries is omnipresent. The trading industry had a skewness of 36.4. On the contrary, skewness in the candy & soda industry was only 4.4. Varying degrees of skewness have implications for investors whom attempt stock picking. Some industries with high skewness and low median lifetime buy-and-hold returns could be seen as "lottery industries" as most stocks in these industries have a negative lifetime return and only a few stocks perform exceptionally well.

Lifetime returns are heavily influenced by how long a stock is actually listed. Stocks that survive for decades have a far greater chance of posting an exceptional lifetime buy-and-hold return compared to stocks that are only listed for a few years. For example, Altria, a tobacco firm that sells well-known cigarette brands such as Marlboro, has existed throughout the full sample of 93 years. Which allowed the firm to crank out an extraordinary buy-and-hold return of 217.71 million %. It comes as no surprise that the tobacco industry is also the industry with the highest mean return. Moreover, the tobacco industry had the highest median return and tobacco firms had the longest life on average compared to other industries. Because firms switch industries and delist quite often, the mean time in the sample was relatively low at 109 months or roughly 9 year.

How powerful compounding is can also be observed from Table 28 in Appendix A, because it covers a much shorter period. The mean lifetime buy-and-hold return in the period from 1926 to 2018 is 5,086% whereas the mean is 617% in the period from 1970 to 2018. The median lifetime buy-and-hold return of stocks is -1% in the period from 1970 to 2018. Therefore, this paper accepts hypothesis 1: long-term buy-and-hold stock returns across US industries are positively skewed and differences in skewness exist across industries. In fact, evidence is found that skewness in buy-and-hold returns varies widely varies significantly across industries and investment horizons.

Table 4: CRSP stocks buy-and-hold returns at monthly horizon during 1926-2018

| | Monthly buy-and-hold returns | | | | |
|---|---|---|---|---|---|
| | Mean | Median | Standard deviation | Skewness | Obs |
| **Fama-French 49 Industries** | | | | | |
| Agriculture | 0.0071 | 0.0000 | 0.186 | 4.127 | 12,639 |
| Food Products | 0.0113 | 0.0000 | 0.136 | 4.825 | 82,644 |
| Candy & Soda | 0.0130 | 0.0049 | 0.119 | 2.413 | 10,898 |
| Beer & Liquor | 0.0097 | 0.0000 | 0.132 | 2.931 | 17,482 |
| Tobacco Products | 0.0114 | 0.0028 | 0.119 | 3.787 | 13,538 |
| Recreation | 0.0072 | 0.0000 | 0.187 | 3.610 | 33,896 |
| Entertainment | 0.0094 | 0.0000 | 0.215 | 7.048 | 48,003 |
| Printing & Publishing | 0.0112 | 0.0000 | 0.155 | 4.789 | 36,169 |
| Consumer Goods | 0.0105 | 0.0000 | 0.155 | 5.721 | 78,930 |
| Apparel | 0.0088 | 0.0000 | 0.158 | 2.846 | 54,324 |
| Healthcare | 0.0119 | 0.0000 | 0.204 | 4.142 | 54,267 |
| Medical Equipment | 0.0106 | 0.0000 | 0.220 | 18.817 | 82,708 |
| Pharmaceutical Products | 0.0123 | 0.0000 | 0.234 | 7.538 | 134,103 |
| Chemicals | 0.0117 | 0.0000 | 0.142 | 2.524 | 76,170 |
| Rubber and Plastic Products | 0.0113 | 0.0000 | 0.165 | 2.138 | 29,169 |
| Textiles | 0.0091 | 0.0000 | 0.148 | 2.179 | 39,744 |
| Construction Materials | 0.0123 | 0.0000 | 0.151 | 3.673 | 112,071 |
| Construction | 0.0093 | 0.0000 | 0.192 | 4.990 | 40,165 |
| Steel Works | 0.0119 | 0.0000 | 0.161 | 22.140 | 82,203 |
| Fabricated Products | 0.0107 | 0.0000 | 0.167 | 4.968 | 13,446 |
| Machinery | 0.0123 | 0.0000 | 0.157 | 4.668 | 132,358 |
| Electrical Equipment | 0.0113 | 0.0000 | 0.208 | 5.803 | 87,520 |
| Automobiles and Trucks | 0.0102 | 0.0000 | 0.151 | 2.465 | 71,673 |
| Aircraft | 0.0140 | 0.0000 | 0.141 | 2.712 | 25,407 |
| Shipbuilding, Railroad Equipment | 0.0125 | 0.0000 | 0.146 | 2.499 | 12,618 |
| Defense | 0.0135 | 0.0022 | 0.151 | 1.480 | 5,606 |
| Precious Metals | 0.0073 | 0.0000 | 0.216 | 2.978 | 36,682 |
| Mines | 0.0085 | 0.0000 | 0.185 | 5.685 | 37,180 |
| Coal | 0.0117 | 0.0000 | 0.208 | 6.244 | 11,288 |
| Petroleum and Natural Gas | 0.0086 | 0.0000 | 0.195 | 8.216 | 180,573 |
| Utilities | 0.0106 | 0.0080 | 0.085 | 4.287 | 128,811 |
| Communication | 0.0101 | 0.0000 | 0.213 | 5.985 | 81,069 |
| Personal Services | 0.0087 | 0.0000 | 0.198 | 8.065 | 34,575 |
| Business Services | 0.0106 | 0.0000 | 0.202 | 4.089 | 185,549 |
| Computers | 0.0113 | 0.0000 | 0.215 | 3.513 | 88,591 |
| Computer Software | 0.0117 | 0.0000 | 0.252 | 5.143 | 143,149 |
| Electronic Equipment | 0.0131 | 0.0000 | 0.203 | 5.090 | 161,641 |
| Measuring and Control Equipment | 0.0139 | 0.0000 | 0.198 | 9.959 | 63,753 |
| Business Supplies | 0.0113 | 0.0000 | 0.144 | 7.769 | 35,003 |
| Shipping Containers | 0.0132 | 0.0025 | 0.139 | 5.637 | 27,565 |
| Transportation | 0.0101 | 0.0000 | 0.169 | 3.825 | 116,806 |
| Wholesale | 0.0099 | 0.0000 | 0.192 | 5.058 | 129,035 |
| Retail | 0.0098 | 0.0000 | 0.166 | 4.392 | 193,126 |
| Restaurant, Hotels, Motels | 0.0083 | 0.0000 | 0.164 | 3.372 | 70,724 |
| Banking | 0.0097 | 0.0034 | 0.126 | 4.310 | 249,167 |
| Insurance | 0.0110 | 0.0025 | 0.129 | 4.357 | 94,533 |
| Real Estate | 0.0078 | 0.0000 | 0.143 | 7.127 | 94,741 |
| Trading | 0.0125 | 0.0000 | 0.137 | 3.981 | 233,353 |
| Other | 0.0088 | 0.0000 | 0.197 | 3.746 | 12,586 |
| Not available | 0.0047 | 0.0000 | 0.197 | 2.982 | 19,987 |
| Total | 0.0107 | 0.0000 | 0.177 | 6.439 | 3,817,238 |

All common stocks and real estate investment trusts (REITs), occurring at some point in time, in CRSP from January 1926 until December 2018 are included. Stocks are assigned to an industry based on their Standard Industry Classification (SIC) code. Kenneth French's data library is used to structure which SIC codes belong to which industry. Companies are allowed to switch industries multiple times. If companies switch back to an industry they were previously in, this is treated as a separate run of returns to prevent faulty linking of returns. Mean and median returns depicted in the table are monthly buy-and-hold returns including dividends and delisting returns.

Table 5: CRSP stocks buy-and-hold returns at annual horizon grouped by industry during 1926-2018

| | Annual buy-and-hold returns | | | | |
|---|---|---|---|---|---|
| | Mean | Median | Standard deviation | Skewness | Obs |
| **Fama-French 49 Industries** | | | | | |
| Agriculture | 0.0738 | 0.0000 | 0.624 | 4.227 | 1,174 |
| Food Products | 0.1377 | 0.0749 | 0.538 | 5.429 | 7,275 |
| Candy & Soda | 0.1778 | 0.0908 | 0.556 | 3.393 | 972 |
| Beer & Liquor | 0.1315 | 0.0602 | 0.574 | 5.603 | 1,542 |
| Tobacco Products | 0.1489 | 0.0986 | 0.500 | 4.377 | 1,163 |
| Recreation | 0.1117 | 0.0000 | 0.935 | 11.985 | 3,100 |
| Entertainment | 0.1228 | 0.0000 | 0.877 | 6.309 | 4,471 |
| Printing & Publishing | 0.1392 | 0.0781 | 0.674 | 11.886 | 3,234 |
| Consumer Goods | 0.1382 | 0.0500 | 0.707 | 10.479 | 7,058 |
| Apparel | 0.1168 | 0.0370 | 0.621 | 3.891 | 4,824 |
| Healthcare | 0.1679 | 0.0000 | 1.060 | 11.590 | 5,116 |
| Medical Equipment | 0.1317 | 0.0000 | 0.917 | 8.062 | 7,582 |
| Pharmaceutical Products | 0.1481 | -0.0014 | 0.981 | 5.484 | 12,182 |
| Chemicals | 0.1433 | 0.0762 | 0.554 | 3.244 | 6,750 |
| Rubber and Plastic Products | 0.1394 | 0.0385 | 0.666 | 3.892 | 2,664 |
| Textiles | 0.1313 | 0.0386 | 0.623 | 2.904 | 3,520 |
| Construction Materials | 0.1494 | 0.0594 | 0.597 | 3.522 | 9,950 |
| Construction | 0.1171 | 0.0028 | 0.682 | 2.557 | 3,682 |
| Steel Works | 0.1410 | 0.0608 | 0.564 | 2.982 | 7,192 |
| Fabricated Products | 0.1202 | 0.0376 | 0.569 | 1.996 | 1,230 |
| Machinery | 0.1500 | 0.0643 | 0.630 | 5.068 | 11,747 |
| Electrical Equipment | 0.1532 | 0.0000 | 1.039 | 10.163 | 8,026 |
| Automobiles and Trucks | 0.1438 | 0.0455 | 0.667 | 4.235 | 6,295 |
| Aircraft | 0.1760 | 0.0910 | 0.574 | 2.800 | 2,234 |
| Shipbuilding, Railroad Equipment | 0.1430 | 0.0695 | 0.545 | 3.396 | 1,099 |
| Defense | 0.2085 | 0.0804 | 1.101 | 11.556 | 505 |
| Precious Metals | 0.0965 | -0.0455 | 1.183 | 22.228 | 3,376 |
| Mines | 0.0987 | 0.0000 | 0.683 | 3.863 | 3,406 |
| Coal | 0.0992 | 0.0245 | 0.547 | 1.049 | 1,004 |
| Petroleum and Natural Gas | 0.1107 | 0.0000 | 0.764 | 5.190 | 16,334 |
| Utilities | 0.1330 | 0.1100 | 0.332 | 3.430 | 11,080 |
| Communication | 0.1591 | 0.0306 | 0.915 | 5.640 | 7,565 |
| Personal Services | 0.1189 | 0.0126 | 0.835 | 9.156 | 3,227 |
| Business Services | 0.1279 | 0.0082 | 0.782 | 5.191 | 17,340 |
| Computers | 0.1427 | -0.0074 | 0.993 | 9.357 | 8,115 |
| Computer Software | 0.1586 | -0.0007 | 1.147 | 8.632 | 13,631 |
| Electronic Equipment | 0.1722 | 0.0000 | 0.941 | 6.702 | 14,599 |
| Measuring and Control Equipment | 0.1721 | 0.0174 | 0.997 | 11.934 | 5,808 |
| Business Supplies | 0.1362 | 0.0749 | 0.554 | 5.549 | 3,127 |
| Shipping Containers | 0.1547 | 0.0851 | 0.521 | 3.308 | 2,436 |
| Transportation | 0.1217 | 0.0448 | 0.639 | 4.923 | 10,332 |
| Wholesale | 0.1257 | 0.0130 | 0.757 | 4.783 | 11,919 |
| Retail | 0.1410 | 0.0387 | 0.710 | 5.046 | 17,323 |
| Restaurant, Hotels, Motels | 0.1292 | 0.0214 | 1.378 | 52.890 | 6,406 |
| Banking | 0.1257 | 0.0789 | 0.479 | 3.747 | 23,095 |
| Insurance | 0.1369 | 0.0899 | 0.504 | 4.306 | 8,445 |
| Real Estate | 0.0961 | 0.0625 | 0.529 | 7.522 | 8,642 |
| Trading | 0.1628 | 0.0893 | 0.938 | 76.170 | 21,805 |
| Other | 0.0911 | 0.0113 | 0.692 | 4.306 | 1,190 |
| Not available | 0.0523 | 0.0000 | 0.677 | 5.151 | 2,151 |
| Total | 0.1377 | 0.0408 | 0.790 | 20.211 | 346,943 |

All common stocks and real estate investment trusts (REITs), occurring at some point in time, in CRSP from January 1926 until December 2018 are included. Stocks are assigned to an industry based on their Standard Industry Classification (SIC) code. Kenneth French's data library is used to structure which SIC codes belong to which industry. Companies are allowed to switch industries multiple times. If companies switch back to an industry they were previously in, this is treated as a separate run of returns to prevent faulty linking of returns. Mean and median returns depicted in the table are annual buy-and-hold returns including dividends and delisting returns.

Table 6: CRSP stocks buy-and-hold returns at decade horizon grouped by industry during 1926-2018

| | **Decade buy-and-hold returns** | | | | **Time in sample** | | |
|---|---|---|---|---|---|---|---|
| | Mean | Median | SD | Skewness | Mean (months) | SD | Obs |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 0.71 | -0.18 | 4.1 | 8.1 | 53 | 40 | 238 |
| Food Products | 1.25 | 0.50 | 2.7 | 5.5 | 74 | 45 | 1,117 |
| Candy & Soda | 1.64 | 0.52 | 3.1 | 2.3 | 68 | 42 | 161 |
| Beer & Liquor | 1.48 | 0.43 | 6.2 | 12.1 | 71 | 43 | 247 |
| Tobacco Products | 1.93 | 0.88 | 3.2 | 2.4 | 94 | 47 | 144 |
| Recreation | 0.82 | -0.26 | 6.6 | 15.0 | 59 | 40 | 578 |
| Entertainment | 0.74 | -0.17 | 3.4 | 9.0 | 52 | 38 | 924 |
| Printing & Publishing | 1.47 | 0.26 | 4.0 | 5.2 | 68 | 42 | 531 |
| Consumer Goods | 1.40 | 0.12 | 5.5 | 12.4 | 67 | 42 | 1,178 |
| Apparel | 0.98 | 0.09 | 2.9 | 4.7 | 71 | 42 | 767 |
| Healthcare | 1.07 | -0.09 | 5.3 | 12.0 | 49 | 36 | 1,103 |
| Medical Equipment | 0.77 | -0.09 | 3.1 | 5.7 | 57 | 40 | 1,457 |
| Pharmaceutical Products | 1.08 | -0.18 | 8.4 | 28.7 | 59 | 41 | 2,271 |
| Chemicals | 1.36 | 0.41 | 3.2 | 5.4 | 71 | 44 | 1,077 |
| Rubber and Plastic Products | 1.05 | 0.11 | 3.3 | 5.3 | 59 | 40 | 496 |
| Textiles | 1.07 | 0.08 | 2.9 | 3.4 | 70 | 44 | 567 |
| Construction Materials | 1.18 | 0.37 | 2.8 | 4.2 | 70 | 42 | 1,606 |
| Construction | 0.66 | -0.09 | 3.0 | 9.0 | 55 | 41 | 726 |
| Steel Works | 1.19 | 0.36 | 2.9 | 5.0 | 78 | 44 | 1,049 |
| Fabricated Products | 0.84 | 0.16 | 3.2 | 9.1 | 58 | 41 | 230 |
| Machinery | 1.15 | 0.32 | 3.1 | 7.6 | 69 | 43 | 1,929 |
| Electrical Equipment | 0.84 | -0.10 | 5.0 | 20.0 | 55 | 40 | 1,596 |
| Automobiles and Trucks | 1.23 | 0.25 | 3.6 | 7.0 | 76 | 44 | 941 |
| Aircraft | 1.72 | 0.55 | 3.4 | 2.9 | 74 | 42 | 345 |
| Shipbuilding, Railroad Equipment | 1.28 | 0.30 | 3.0 | 4.2 | 79 | 48 | 159 |
| Defense | 1.34 | 0.39 | 2.9 | 2.8 | 60 | 42 | 94 |
| Precious Metals | 0.29 | -0.46 | 2.6 | 6.0 | 57 | 40 | 643 |
| Mines | 0.71 | -0.10 | 2.9 | 5.3 | 61 | 43 | 611 |
| Coal | 0.95 | 0.13 | 2.8 | 4.1 | 68 | 47 | 167 |
| Petroleum and Natural Gas | 0.73 | -0.07 | 2.8 | 5.3 | 65 | 41 | 2,760 |
| Utilities | 1.61 | 0.96 | 2.0 | 1.7 | 88 | 42 | 1,464 |
| Communication | 1.19 | 0.05 | 6.6 | 16.4 | 51 | 39 | 1,576 |
| Personal Services | 0.87 | -0.12 | 4.4 | 10.6 | 51 | 37 | 677 |
| Business Services | 0.83 | -0.04 | 4.2 | 13.4 | 51 | 40 | 3,621 |
| Computers | 1.03 | -0.21 | 11.7 | 25.7 | 55 | 40 | 1,598 |
| Computer Software | 0.72 | -0.32 | 7.1 | 34.9 | 48 | 36 | 2,995 |
| Electronic Equipment | 1.03 | -0.02 | 3.9 | 6.9 | 63 | 41 | 2,571 |
| Measuring and Control Equipment | 0.94 | 0.08 | 3.4 | 6.8 | 61 | 41 | 1,050 |
| Business Supplies | 1.24 | 0.46 | 3.1 | 5.8 | 66 | 41 | 533 |
| Shipping Containers | 1.54 | 0.45 | 3.0 | 2.7 | 70 | 42 | 394 |
| Transportation | 1.11 | 0.14 | 4.3 | 16.4 | 71 | 45 | 1,644 |
| Wholesale | 0.80 | -0.03 | 4.1 | 16.6 | 53 | 40 | 2,428 |
| Retail | 1.21 | 0.11 | 4.1 | 7.2 | 66 | 42 | 2,947 |
| Restaurant, Hotels, Motels | 0.96 | -0.02 | 4.4 | 12.2 | 62 | 40 | 1,148 |
| Banking | 0.95 | 0.32 | 2.7 | 11.5 | 55 | 39 | 4,553 |
| Insurance | 1.22 | 0.36 | 3.1 | 6.5 | 64 | 41 | 1,476 |
| Real Estate | 0.74 | 0.13 | 2.4 | 6.9 | 58 | 39 | 1,641 |
| Trading | 1.24 | 0.20 | 3.9 | 13.5 | 48 | 38 | 4,820 |
| Other | 0.82 | -0.11 | 3.8 | 6.4 | 49 | 40 | 258 |
| Not available | 0.21 | 0.00 | 2.0 | 8.8 | 29 | 33 | 685 |
| Total | 1.02 | 0.09 | 4.6 | 29.7 | 60 | 42 | 63,791 |

All common stocks and real estate investment trusts (REITs), occurring at some point in time, in CRSP from January 1926 until December 2018 are included. Stocks are assigned to an industry based on their Standard Industry Classification (SIC) code. Kenneth French's data library is used to structure which SIC codes belong to which industry. Companies are allowed to switch industries multiple times. If companies switch back to an industry they were previously in, this is treated as a separate run of returns to prevent faulty linking of returns. Mean and median returns depicted in the table are decade buy-and-hold returns including dividends and delisting returns.

Table 7: CRSP stocks buy-and-hold returns at lifetime horizon grouped by industry during 1926-2018

| | **Lifetime buy-and-hold returns** | | | | **Time in sample** | | |
|---|---|---|---|---|---|---|---|
| | Mean | Median | SD | Skewness | Mean (months) | SD | Obs |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 1.89 | -0.31 | 8.6 | 6.7 | 92 | 97 | 138 |
| Food Products | 115.81 | 0.60 | 1305.0 | 17.4 | 167 | 192 | 494 |
| Candy & Soda | 55.47 | 0.49 | 214.9 | 4.4 | 135 | 151 | 81 |
| Beer & Liquor | 71.58 | 0.32 | 453.8 | 8.7 | 163 | 180 | 107 |
| Tobacco Products | 3688.82 | 1.09 | 19344.4 | 5.9 | 282 | 288 | 48 |
| Recreation | 4.85 | -0.50 | 48.2 | 15.1 | 104 | 107 | 327 |
| Entertainment | 15.90 | -0.40 | 249.0 | 22.4 | 87 | 94 | 551 |
| Printing & Publishing | 117.94 | 0.15 | 1739.0 | 16.3 | 133 | 145 | 272 |
| Consumer Goods | 44.47 | 0.04 | 529.3 | 16.3 | 139 | 152 | 568 |
| Apparel | 28.58 | -0.12 | 265.5 | 13.6 | 150 | 161 | 361 |
| Healthcare | 3.15 | -0.13 | 25.3 | 15.5 | 78 | 75 | 693 |
| Medical Equipment | 17.02 | -0.28 | 281.7 | 25.0 | 99 | 101 | 832 |
| Pharmaceutical Products | 111.31 | -0.42 | 2026.9 | 23.0 | 99 | 109 | 1,351 |
| Chemicals | 67.06 | 0.60 | 589.6 | 16.0 | 151 | 179 | 506 |
| Rubber and Plastic Products | 6.82 | 0.11 | 40.0 | 9.2 | 108 | 105 | 270 |
| Textiles | 11.55 | 0.16 | 82.7 | 12.4 | 155 | 154 | 256 |
| Construction Materials | 32.67 | 0.52 | 464.5 | 21.0 | 150 | 154 | 745 |
| Construction | 3.15 | -0.17 | 27.8 | 13.6 | 96 | 103 | 418 |
| Steel Works | 36.18 | 0.47 | 286.8 | 13.4 | 193 | 200 | 426 |
| Fabricated Products | 3.18 | 0.08 | 11.1 | 5.1 | 108 | 93 | 125 |
| Machinery | 122.34 | 0.32 | 1526.8 | 17.1 | 146 | 168 | 909 |
| Electrical Equipment | 16.76 | -0.34 | 383.3 | 29.2 | 102 | 104 | 862 |
| Automobiles and Trucks | 38.28 | 0.09 | 319.9 | 16.2 | 172 | 186 | 417 |
| Aircraft | 2247.42 | 0.77 | 26627.5 | 12.1 | 169 | 182 | 150 |
| Shipbuilding, Railroad Equipment | 26.42 | 0.30 | 127.0 | 7.1 | 183 | 213 | 69 |
| Defense | 18.29 | 0.28 | 96.1 | 6.8 | 106 | 114 | 53 |
| Precious Metals | 0.86 | -0.70 | 7.2 | 8.6 | 102 | 96 | 359 |
| Mines | 12.21 | -0.33 | 140.9 | 16.7 | 106 | 124 | 350 |
| Coal | 4.36 | 0.17 | 13.3 | 3.6 | 136 | 136 | 83 |
| Petroleum and Natural Gas | 54.07 | -0.48 | 732.6 | 22.1 | 117 | 140 | 1,544 |
| Utilities | 191.43 | 3.36 | 636.3 | 5.5 | 279 | 276 | 461 |
| Communication | 8.52 | -0.14 | 79.9 | 17.5 | 86 | 95 | 943 |
| Personal Services | 2.46 | -0.20 | 21.8 | 17.4 | 85 | 86 | 409 |
| Business Services | 7.78 | -0.16 | 196.9 | 44.6 | 83 | 91 | 2,230 |
| Computers | 148.45 | -0.44 | 4242.8 | 29.6 | 101 | 106 | 877 |
| Computer Software | 4.22 | -0.50 | 56.8 | 21.2 | 73 | 71 | 1,955 |
| Electronic Equipment | 6.32 | -0.22 | 60.4 | 21.7 | 118 | 115 | 1,369 |
| Measuring and Control Equipment | 4.01 | 0.03 | 19.5 | 8.8 | 111 | 109 | 576 |
| Business Supplies | 10.46 | 0.59 | 48.3 | 9.6 | 132 | 132 | 265 |
| Shipping Containers | 22.54 | 0.82 | 151.9 | 11.8 | 152 | 153 | 181 |
| Transportation | 153.36 | -0.01 | 4005.8 | 27.9 | 149 | 158 | 785 |
| Wholesale | 37.67 | -0.16 | 1197.8 | 37.3 | 91 | 103 | 1,411 |
| Retail | 79.22 | -0.06 | 1672.2 | 34.1 | 130 | 145 | 1,487 |
| Restaurant, Hotels, Motels | 34.70 | -0.21 | 598.4 | 20.9 | 112 | 118 | 634 |
| Banking | 3.62 | 0.65 | 55.7 | 49.0 | 90 | 84 | 2,771 |
| Insurance | 9.82 | 0.62 | 73.5 | 20.1 | 123 | 116 | 767 |
| Real Estate | 2.25 | 0.15 | 8.6 | 7.1 | 98 | 98 | 964 |
| Trading | 5.27 | 0.53 | 51.8 | 36.4 | 83 | 83 | 2,796 |
| Other | 1.42 | -0.24 | 5.7 | 4.7 | 80 | 81 | 158 |
| Not available | 0.89 | 0.00 | 11.2 | 18.4 | 35 | 57 | 564 |
| Total | 50.36 | 0.02 | 2209.8 | 108.5 | 109 | 126 | 34,968 |

All common stocks and real estate investment trusts (REITs), occurring at some point in time, in CRSP from January 1926 until December 2018 are included. Stocks are assigned to an industry based on their Standard Industry Classification (SIC) code. Kenneth French's data library is used to structure which SIC codes belong to which industry. Companies are allowed to switch industries multiple times. If companies switch back to an industry they were previously in, this is treated as a separate run of returns to prevent faulty linking of returns. Mean and median returns depicted in the table are lifetime buy-and-hold returns including dividends and delisting returns.

### 4.4 Treasury bills and inflation

Hypotheses 2, 3, 4 and 5 make claims about the ability of stocks to achieve a higher buy-and-hold return than the buy-and-hold return of risk-free asset. This paper uses US Treasury bills as the risk-free asset. Monthly US Treasury bill returns are retrieved from CRSP. To calculate long-term return series, monthly US Treasury bill returns are compounded following equation (7). Hypotheses 6, 7, 8 and 9 make statements about the concentration of wealth creation. Wealth creation by stocks is defined as: "the difference between the present US dollar value of investors' actual investments in the stock and the value that would have been obtained if the same capital investments had earned *one-month US Treasury bill returns*". Hence, monthly US Treasury bills from CRSP are used.

In the robustness section of this paper an alternative method for calculating wealth creation by stocks is proposed. This method is based on the rate of inflation instead of the return of the risk-free asset. In the robustness section, all hypotheses are tested again using the alternative methods. This paper uses monthly increases or decreases in the consumer price index (CPI) as the monthly rate of inflation. CPI data is retrieved from CRSP.

### 4.5 Corporate performance

Corporate performance and the variability in corporate performance are used to test hypotheses 3-5 and 7-9. In those hypotheses, this paper infers that corporate performance and its variability could explain cross-industry variation the proportion of stocks that outperform the risk-free rate over their lifetime and cross-industry variation in concentration of wealth creation. Therefore, in the hypothesis development section this paper identifies five categories of corporate performance to test the claims of hypotheses 3-5 and 7-9. The five categories of corporate performance are capitalization, solvency, liquidity, efficiency and profitability. The categories provide a broad overview of the firms' corporate performance and cover the balance sheet, profit & loss account and cash flows of a firm. Table 8 shows which financial ratios are used to measure corporate performance. The financial ratios are selected based on data availability, common acceptance by financial practitioners and appearance in financial economics academic literature.

The ratios are extracted from the Financial Ratios Suite of WRDS, which provides ratios on an industry level. WRDS starts out with a large sample of quarterly and annual accounting data from Compustat which can be linked to firms from the CRSP database. They subsequently populate the data to monthly frequency, by carrying forward the most recent quarterly or annual data. WRDS also imposes a two month lag to all observations to avoid a look ahead bias. They acknowledge that ratio metrics can produce extreme outliers. Therefore, WRDS imposes winsorization and smoothing before aggregating ratios on the industry level. WRDS conveniently compiles industry ratios based on the Fama-French industry classification, which this paper also uses. To arrive at the industry level WRDS takes the median of all the firms within that industry for every month[5]. Financial ratios are available in the period from 1970 to 2018. Therefore, subsequent sections always show statistics for 1926-2018 and 1970-2018.

[5] Median instead of mean ratios are used, because the denominator of some ratios can be negative resulting

Table 8: Corporate performance ratios

| Category | Financial ratio | Formula |
|---|---|---|
| Capitalization | Capitalization ratio | Total long-term debt as a fraction of the sum of total long-term debt, common/ordinary equity and preferred stock. |
| Solvency | Cash flow/total debt | Operating cash flow as a fraction of total debt. |
| Solvency | Interest coverage | Multiple of earnings before interest and taxes to interest and related expenses. |
| Liquidity | Quick ratio | Current assets net of inventories as a fraction of current liabilities |
| Efficiency | Asset turnover | Sales as a fraction of the average total assets based on the most recent two periods. |
| Profitability | Return on capital employed (ROCE) | Earnings before interest and taxes as a fraction of average capital employed based on most recent two periods. Where capital employed is the sum of debt in long-term and current liabilities and common/ordinary equity. |
| Profitability | Operating margin | Operating income after depreciation as a fraction of sales. |

Table 9 reports the average ratios over the length of the samples per industry, these are used as the inputs for later regressions. Accounting data is unavailable for the ”Not available” industry class, because WRDS excludes those firms. They represent a very small portion of the total sample which is why we are of the opinion that missing accounting data for this group of stocks is not problematic.

in a meaningless ratio.

Table 9: Average of median financial ratios per industry during 1970-2018

| | Average median financial ratios per industry | | | | | | |
|---|---|---|---|---|---|---|---|
| | **Capitalization** | **Efficiency** | **Liquidity** | **Profitability** | | **Solvency** | |
| | Capital ratio | Asset turnover | Quick ratio | Return on assets | Operating margin | Cash flow/debt | Interest coverage |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 0.31 | 0.87 | 1.16 | 8.9% | 5.8% | 0.09 | 3.11 |
| Food Products | 0.29 | 1.65 | 1.01 | 14.3% | 6.6% | 0.16 | 4.97 |
| Candy & Soda | 0.37 | 1.29 | 1.09 | 15.5% | 8.1% | 0.17 | 5.19 |
| Beer & Liquor | 0.25 | 1.03 | 1.01 | 13.2% | 10.6% | 0.17 | 5.09 |
| Tobacco Products | 0.46 | 0.96 | 0.86 | 18.2% | 21.7% | 0.16 | 7.56 |
| Recreation | 0.15 | 1.34 | 1.40 | 11.2% | 5.6% | 0.11 | 3.49 |
| Entertainment | 0.38 | 0.77 | 1.09 | 12.3% | 8.3% | 0.13 | 2.07 |
| Printing & Publishing | 0.26 | 1.02 | 1.28 | 14.9% | 10.8% | 0.17 | 6.36 |
| Consumer Goods | 0.22 | 1.39 | 1.29 | 14.0% | 7.7% | 0.15 | 6.29 |
| Apparel | 0.15 | 1.63 | 1.45 | 14.0% | 7.0% | 0.16 | 7.01 |
| Healthcare | 0.36 | 1.04 | 1.65 | 13.5% | 8.2% | 0.10 | 3.03 |
| Medical Equipment | 0.08 | 1.02 | 2.14 | 7.9% | 3.9% | 0.03 | 3.65 |
| Pharmaceutical Products | 0.07 | 0.65 | 3.22 | -9.5% | -49.6% | -0.34 | -4.73 |
| Chemicals | 0.30 | 1.08 | 1.32 | 14.6% | 9.2% | 0.16 | 5.03 |
| Rubber and Plastic Products | 0.30 | 1.36 | 1.31 | 13.9% | 7.0% | 0.15 | 4.12 |
| Textiles | 0.33 | 1.42 | 1.39 | 13.3% | 6.7% | 0.14 | 4.13 |
| Construction Materials | 0.27 | 1.22 | 1.37 | 13.1% | 7.2% | 0.15 | 4.57 |
| Construction | 0.36 | 1.27 | 1.29 | 9.1% | 5.0% | 0.06 | 2.64 |
| Steel Works | 0.30 | 1.23 | 1.22 | 11.4% | 5.8% | 0.12 | 3.73 |
| Fabricated Products | 0.24 | 1.24 | 1.29 | 11.9% | 6.9% | 0.13 | 7.89 |
| Machinery | 0.22 | 1.16 | 1.33 | 12.4% | 7.6% | 0.13 | 5.48 |
| Electrical Equipment | 0.17 | 1.22 | 1.42 | 12.7% | 6.9% | 0.12 | 5.42 |
| Automobiles and Trucks | 0.29 | 1.45 | 1.14 | 13.0% | 6.0% | 0.12 | 4.22 |
| Aircraft | 0.30 | 1.20 | 1.03 | 13.1% | 8.3% | 0.12 | 5.51 |
| Shipbuilding, Railroad Equipment | 0.33 | 1.23 | 1.25 | 11.9% | 7.0% | 0.11 | 4.19 |
| Defense | 0.24 | 1.18 | 1.42 | 15.4% | 9.4% | 0.18 | 6.79 |
| Precious Metals | 0.07 | 0.39 | 2.36 | 2.0% | -3.5% | 0.16 | -1.61 |
| Mines | 0.19 | 0.58 | 1.51 | 8.5% | 9.1% | 0.15 | 3.11 |
| Coal | 0.41 | 0.70 | 1.15 | 10.3% | 6.6% | 0.14 | 2.10 |
| Petroleum and Natural Gas | 0.30 | 0.42 | 1.15 | 13.1% | 6.4% | 0.20 | 2.27 |
| Utilities | 0.50 | 0.41 | 0.70 | 11.1% | 18.7% | 0.11 | 2.84 |
| Communication | 0.46 | 0.48 | 1.26 | 12.4% | 12.8% | 0.12 | 1.83 |
| Personal Services | 0.29 | 1.12 | 1.27 | 13.7% | 8.1% | 0.15 | 3.27 |
| Business Services | 0.16 | 1.18 | 1.62 | 12.7% | 6.3% | 0.13 | 4.04 |
| Computers | 0.10 | 1.10 | 1.75 | 9.9% | 4.1% | 0.09 | 2.83 |
| Computer Software | 0.08 | 1.02 | 1.96 | 10.0% | 2.9% | 0.13 | 2.27 |
| Electronic Equipment | 0.12 | 1.09 | 1.91 | 11.0% | 5.2% | 0.14 | 3.80 |
| Measuring and Control Equipment | 0.10 | 1.05 | 1.96 | 11.3% | 6.9% | 0.13 | 5.91 |
| Business Supplies | 0.32 | 1.27 | 1.23 | 15.0% | 8.1% | 0.17 | 4.96 |
| Shipping Containers | 0.44 | 1.29 | 0.98 | 15.3% | 8.3% | 0.15 | 4.31 |
| Transportation | 0.41 | 1.18 | 1.18 | 13.1% | 6.3% | 0.15 | 2.98 |
| Wholesale | 0.26 | 2.00 | 1.10 | 10.7% | 3.5% | 0.07 | 4.04 |
| Retail | 0.24 | 2.22 | 0.62 | 14.0% | 4.1% | 0.14 | 4.99 |
| Restaurant, Hotels, Motels | 0.36 | 1.43 | 0.78 | 14.9% | 6.8% | 0.19 | 3.10 |
| Banking | 0.37 | 0.08 | 1.65 | 2.2% | 26.4% | 0.01 | 1.95 |
| Insurance | 0.13 | 0.38 | 1.43 | 4.6% | 11.2% | 0.07 | 8.99 |
| Real Estate | 0.42 | 0.32 | 1.25 | 5.6% | 9.2% | 0.05 | 1.33 |
| Trading | 0.22 | 0.22 | 1.90 | 4.5% | 22.0% | 0.05 | 3.31 |
| Other | 0.37 | 0.84 | 1.36 | 15.2% | 9.9% | 0.13 | 2.82 |
| Not available | | | | | | | |
| Total | 0.27 | 1.05 | 1.38 | 11.3% | 7.1% | 0.12 | 3.92 |

The table displays the average median financial ratios for every industry during 1970-2018.

# 5 Methodology

This sections outlines how hypotheses 2 to 9 are formally tested. First, calculations for stock return performance against benchmarks such as the risk-free asset are formulated. Second, wealth creation and concentration of wealth creation calculations are specified. Third, independent variables used in the regression analysis are discussed. Fourth, the decision to use fractional outcome regressions instead of ordinary least squares regressions is motivated. Fifth, equations for the fractional outcome regression used in this paper are provided. Last, marginal effects plots are explained which help to intuitively understand results from fractional outcome regressions. Because a large number of hypotheses are tested, Table 10 below provides a summary of how hypotheses 2 to 9 are formally tested.

## 5.1 Benchmarking stocks' performance across US industries

Hypotheses 2, 3, 4 and 5 all make claims about the ability of stocks to outperform the risk-free rate over their lifetime. To correctly determine whether an individual stock outperformed US Treasury bills, times frames over which returns are calculated must match exactly. As explained in section 4.2.1, firms behind stocks switch industries frequently. Therefore, panels based on industry are created within individual stock panels. To measure whether a stock outperformed US Treasury bills, buy-and-hold return of both are calculated simultaneously. Subsequently, a dummy indicating either 0 or 1 is created for every panel to indicate whether the stock outperformed US Treasury bills or not.

Taking the mean of the dummy variable for every industry produces proportions indicating how many stocks outperformed US Treasury bills. These can be used to test hypothesis 2. Hence, we perform a t-test using those proportions to establish whether the majority of stocks across US industries are unable to achieve a higher buy-and-hold return over their lifetime than the return of US Treasury bills. Proportions that indicate how many stocks outperformed US Treasury bills over their lifetime are also used as the dependent variable when testing hypotheses 3, 4 and 5.

In addition to measuring whether stocks are able to outperform US Treasury bills, performance against six other benchmarks is calculated to provide a more detailed overview of how stocks perform in the long-term. Which brings the total number of benchmarks to seven:

1. Return of US Treasury bills
2. Return of 0%
3. Rate of inflation (CPI)
4. Equal-weighted market return
5. Value-weighted market return
6. Equal-weighted industry return
7. Value-weighted industry return

Table 10: Hypothesis testing overview

| Hypothesis | Time period | Formal test | Dependent variables | Independent variables |
|---|---|---|---|---|
| H2: the majority (>50%) of stocks across US industries are unable to achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset (US Treasury bills). | 1926-2018<br>1970-2018 | T-test | - | - |
| H3: short-term stock return volatility, return skewness and mean returns can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset. | 1926-2018<br>1970-2018 | Fractional outcome regression | Proportion of stocks outperforming T-bills over their lifetime. | Short-term stock return volatility, short-term return skewness and short-term mean returns. |
| H4: corporate performance can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the return of the risk-free rate. | 1970-2018 | Fractional outcome regression | Proportion of stocks outperforming T-bills over their lifetime. | Median industry capitalization ratio, asset turnover, quick ratio, ROCE, operating margin, cash flow/debt and interest coverage. |
| H5: variability in corporate performance can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the risk-free asset. | 1970-2018 | Fractional outcome regression | Proportion of stocks outperforming T-bills over their lifetime. | Average standard deviation of capitalization ratio, asset turnover, quick ratio, ROCE, operating margin, cash/flow debt and interest coverage. |
| H6: a high degree of concentration in wealth creation, i.e. <%5 of stocks are responsible for 100% of wealth creation, is present across US industries. | 1926-2018<br>1970-2018 | T-test | - | - |
| H7: short-term stock return volatility, return skewness and mean returns can explain differences in the proportion of stocks that is responsible for a x% of wealth across US industries. | 1926-2018<br>1970-2018 | Fractional outcome regression | Proportion of gross wealth created by the top 2.5% of stocks.<br>Proportion of stocks responsible for 100% of net wealth creation. | Short-term stock return volatility, short-term return skewness and short-term mean returns. |
| H8: corporate performance can explain differences in the proportion of stocks that is responsible for a x% of wealth across US industries. | 1970-2018 | Fractional outcome regression | Proportion of gross wealth created by the top 2.5% of stocks.<br>Proportion of stocks responsible for 100% of net wealth creation. | Median industry capitalization ratio, asset turnover, quick ratio, ROCE, operating margin, cash flow/debt and interest coverage. |
| H9: variability in corporate performance can explain differences in the proportion of stocks that is responsible for a x% of wealth across US industries. | 1970-2018 | Fractional outcome regression | Proportion of gross wealth created by top 2.5% of stocks.<br>Proportion of stocks responsible for 100% of net wealth creation. | Average standard deviation of capitalization ratio, asset turnover, quick ratio, ROCE, operating margin, cash/flow debt and interest coverage. |

### 5.1.1 Market returns

Measuring market returns could reinforce evidence for hypothesis 1 which conjectures that long-term buy-and-hold stock returns across US industries are positively skewed and differences in skewness exist across industries. Measuring what proportion of stocks in an industry is able to outperform the matched market return gives an indication whether the median return is lower than the mean return. A lower median than mean indicates positive skewness. Both equal-weighted and value-weighted market returns are calculated on a monthly basis. Equal-weighted returns are calculated by taking the mean return across all stocks in the sample. Value-weighted returns are calculated by attaching weights, in this case market capitalization, to every observation and taking the mean. Market capitalization is calculated by taking the product of shares outstanding and price at the beginning of each month. Stocks with missing shares outstanding data for all observations are excluded. To calculate market returns over annual, decade and lifetime horizons the same compounding method from equation (7) is used.

### 5.1.2 Industry returns

Measuring industry returns could also reinforce evidence for hypothesis 1. Which conjectures that long-horizon buy-and-hold stock returns across US industries are positively skewed and differences in skewness exist across industries. Measuring what proportion of stocks in an industry is able to outperform that industries' return directly shows whether the median stock return is lower than the mean stock return. When the median stock return is lower than the mean, the return distribution is positively skewed. Equal-weighted and value-weighted industry returns are calculated in the same fashion as market returns described in 5.1.1 with one extra step. Instead of using a stock's PERMNO as the unique identifier, industry panel IDs are used to identify a continuous run of returns. Going back to the example in Table 3. The panel ID in the last column would be used as the identifier for a return series instead of PERMNO.

## 5.2 Calculating wealth creation by individual stocks

To test hypotheses 6, 7, 8 and 9 concentration of wealth creation by stocks in their industries is calculated. Hypothesis 6 conjectures that a high degree of concentration in wealth creation is present across US industries. Hypotheses 7, 8 and 9 make the inference that short-term stock return volatility, return skewness, mean returns, corporate performance and variation in corporate performance can explain differences in concentration of wealth creation across US industries. To calculate the degree of concentration in wealth creation, wealth creation by each stock must be calculated first. This section describes the method used to calculate wealth creation by stocks.

The reason for calculating wealth creation is that it closely matches the actual experience of investors in aggregate. Buy-and-hold returns with dividends reinvested simulates the experience of an investor buying once and not transacting at all. However, investors in aggregate receive the proceeds of share repurchases, fund equity issuance and do not reinvest dividends. In order to calculate wealth creation by individual stocks, companies, industries and the market as a whole the approach taken by Bessembinder (2018) is followed. The next section of this paper repeats the methods used.

An investor has initial wealth $W_0$ and invests over a time horizon of T periods. The investor has two options to allocate his wealth. First, invest in a riskless bond that pays a fixed rate of return $R_{ft}$ every period $t$. Second, invest in risky assets, i.e. stocks, that have an uncertain return consisting of capital gains and dividends which is expressed as: $R_t$=$R_{ct}$+$R_{dt}$. Received dividends are returned to the investor's risk-free bond account, in this case US Treasury bills.

In every period $t$ the investor has the option to make an additional investment from the risk-free bond account into the risky asset, by the amount of $F_t$. If shares are repurchased by a firm, this is denoted as $F_t < 0$.

The total wealth of the investor at time $t$ is $W_t$. Where the value of the risk-free bond account is $B_t$ and the value of the risky asset is $I_t$:

$$W_t = B_t + I_t \tag{8}$$

The value of the risk-free bond account behaves according to:

$$B_t = B_{t-1} * (1 + R_{ft}) + I_{t-1} * R_{dt} - F_t \tag{9}$$

The first part of the equation is the investor receiving interest from the risk-free bonds. The second part is the investor receiving dividends from the risky asset. The third part describes the investor increasing or decreasing the investment in the risky asset by withdrawing from the risk-free bond account. The value of the risky asset evolves according to:

$$I_t = I_{t-1} * (1 + R_{ct}) + F_t \tag{10}$$

The first part covers the capital gain or loss in the risky asset. The second part covers any net new investment by withdrawing from the risk-free bond account. Adding the two value equations of risk-free asset and the risky asset together yields the overall wealth at time $t$ of the investor:

$$W_t = B_{t-1}(1 + R_f t) + I_{t-1} * (1 + R_t) \tag{11}$$

Which can be rewritten into:

$$W_t - W_{t-1} * (1 + R_{ft}) = I_{t-1} * (R_t - R_f) \tag{12}$$

This expression eliminates dividends, repurchases and new equity investments and shows that they only affect the magnitude of the next period net investment. In essence, the expression states that the excess return over the risk-free rate, $R_t$-$R_f$, times the investment, $I_{t-1}$, is the wealth created above what could be attained by investing only in the risk-free rate.

An investor invests from $t_0$ until $T$. Therefore, if the risky asset has excess return and wealth is created in the past, this wealth must be translated to present values. For example, an investment at *t+1* of \$ 100 with a return of 10% while the risk-free rate was $R_{f+1}$*=2%* creates \$ 8 of wealth. However, at *t+5* this wealth can not be interpreted as \$ 8 anymore, because it could have been invested in the risk-free rate or the risky asset. If it was invested at the risk-free rate and the risk-free rate was constant at 2% the \$8 would be worth *\$8 * 1.02*$^4$ *= \$ 8.66*. Hence, wealth created in previous periods must be adjusted by a future value factor which is the risk-free rate compounded from $t$ to $T$. Such a future value factor, $FV_{t,T}$ can be expressed as:

$$FV_{t,T} = (1 + R_{ft+1}) * (1 + R_{ft+2}) * (1 + R_{ft+3}) * ... * (1 + R_{fT}) \tag{13}$$

There is no need for inflation correction (i.e. translation into "real" terms) as the dollar amounts of wealth created in the past are compounded forward by the risk-free rate $R_{ft}$. By applying the future value factor expression to (12), the present value of wealth created at every period $t$ can be translated into wealth at $T$. Which follows the expression:

$$\begin{aligned} &W_T - W_0 * FV_{0,T} = I_0 * (R_1 - R_{f1}) * FV_{1,T} \\ &+ I_1 * (R_2 - R_{f2}) * FV_{2,T} + ... \\ &+ I_{T-2} * (R_{T-1} - R_{fT-1}) * FV_{T-1,T} \\ &+ I_{T-1} * (R_T - R_{fT}) \end{aligned} \tag{14}$$

Expression (14) can be applied to any stock in the sample by using the market capitalization at the beginning of each period $t$ as the investment $I_t$ in the risky asset. In this paper, the market capitalization at the beginning of every month is used as the investment in the risky asset $I_t$. Monthly buy-and-hold returns including dividends and delisting returns are used as the return $R_t$ of the risky asset. The 30-day Treasury bill rate is used for $R_{ft}$ and for calculating the future value factors $FV_{t,T}$.

Wealth creation among stocks is based on PERMNOs from CRSP which are unique permanent identifiers for a stock. On the firm level this paper aggregates wealth creation by using PERMCOs which are unique permanent identifiers for a firm. The results between a stock and firm only differs if the firm has multiple share classes.

This paper is mainly interested in wealth creation over a stocks' lifetime in an industry. As the use of wealth creation is not widely spread in academic literature it is important to understand how it is calculated beyond the equations described above. By using equation (14) and a made up firm with one share class, Table 11 shows how wealth creation is calculated from monthly observations.

In the second row with date January 2005 one can see that the dollar gain for that month is 4 million dollars, which is the product of excess return and market cap at the beginning of the month: *1.6% * 250 = 4*. To arrive at the wealth gain for that month at present value, the dollar gain is multiplied with the future value factor FV: *4 * 2.19 = 8.76*.

To calculate the total wealth created by a stock over its lifetime in an industry the sum is taken of every months' wealth gain as long as it is in that industry. Hence, for the firm in Table 11, the lifetime wealth created in the banking industry is *8.76 + 13.8975 + ...* until August 2010 when it switches to the insurance industry.

Note how in the row of October 2018 and November 2018 dividends are accounted for. In October 2018 the capital gain of the firm is 2% and a dividend is received of 1%. Therefore, in the absence of share issuance's and repurchases the market cap at the beginning of November 2018 is *700*(1+0.02)=714*. The dollar and wealth gain of October 2018 include the dividend received.

Table 11: Wealth creation example

| PERMNO | Date | Industry | Return | Return T-bill | Excess return | Market cap ($M) | Dollar gain ($M) | FV | Wealth gain ($M) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Jan 2005 | Banking | 2% | 0.4% | 1.6% | 250 | 4 | 2.19 | 8.76 |
| 1 | Feb 2005 | Banking | 3% | 0.5% | 2.5% | 255 | 6.375 | 2.18 | 13.8975 |
| 1 | ... | Banking | ... | ... | ... | ... | ... | ... | ... |
| 1 | Aug 2010 | Insurance | 5% | 0.4% | 4.6% | 400 | 18.4 | 1.06 | 19.504 |
| 1 | ... | Insurance | ... | ... | ... | ... | ... | ... | ... |
| 1 | Oct 2018 | Insurance | 3% (1% dividend) | 0.5% | 2.5% | 700 | 17.5 | 1.010025 | 17.68 |
| 1 | Nov 2018 | Insurance | 4% | 0.5% | 3.5% | 714 | 24.99 | 1.005 | 25.115 |
| 1 | Dec 2018 | Insurance | -2% | 0.5% | -2.5% | 743 | -18.56 | 1 | -18.56 |

In the case of share issuance or repurchases the number of shares outstanding changes. This is reflected in the market capitalization of stocks. In itself a share issuance or repurchase is not a wealth creating or destroying event for investors in aggregate. In the case of a share repurchase the investor selling receives the proceeds of the purchase by the firm, which is available for reinvestment into other risky assets or the risk-free asset. This concludes the methodology used to calculate wealth creation at the stock level.

## 5.3 Aggregating wealth creation to industry and market levels

To calculate concentration of wealth creation one must aggregate wealth creation on an industry or market level. On the industry level, wealth creation is aggregated by using the 49 Fama-French industry categories. Thus, all wealth created by stocks over their lifetime in a certain industry is summed. This paper allows stocks to switch industries during their lifetime to obtain a realistic measurement of industry wealth creation. Stocks switch industries if their primary source of revenue shifts from one industry to another. On the market level, all wealth created by stocks over their lifetime is summed regardless of their industry.

Two distinctions can be made when aggregating wealth creation. First, wealth can be aggregated on a gross basis where only firms are included that have created wealth during their lifetime. Second, wealth can be aggregated on a net basis where all firms are included regardless of whether they destroyed or created wealth during their lifetime.

Gross wealth is defined as the sum of wealth created by companies with positive wealth creation, $Wealth_i > 0$ , where $i$ indicates an individual firm and $N$ is the total number of firms. This is captured in the following equation:

$$\begin{aligned} &Gross\ wealth = \sum_{i=1}^{N} Wealth_i \\ &if\ Wealth_i > 0 \end{aligned} \tag{15}$$

Net wealth creation takes into account all firms, regardless of whether they created or destroyed wealth during their lifetime. Therefore, the if statement from (15) is eliminated:

$$Net\ wealth = \sum_{i=1}^{N} Wealth_i \tag{16}$$

### 5.4 Measuring concentration of gross wealth creation across US industries

This paper is particularly interested in the question if and why wealth creation among industries is concentrated. The previous section describes how wealth creation by stocks is calculated and how aggregation to an industry or market level works. The aggregate figure of total wealth created by an industry or the market is used as the denominator when calculating wealth concentration. The numerator is the amount of wealth created by an individual stock. To answer hypotheses 6, 7, 8 and 9, this paper calculates the degree of concentration in wealth creation.

To calculate the concentration of gross wealth creation, this paper measures what percentage of total gross wealth is created by the top 0.5%, 1% and 2.5% of firms. For example, there are 1000 firms in an industry of which 500 created wealth during their lifetime. This paper measures what percentage of the total gross wealth of that industry is created by the most wealth creating 3 firms (top 0.5%), 5 firms (top 1%) and 13 firms (top 5%). The proportion of gross wealth created by a top $x\%$ of firms can be captured in the following equation:

$$\begin{aligned} &\frac{\sum_{i=1}^{x\%*N} Wealth_i}{\sum_{i=1}^{N} Wealth_i} = Proportion\ gross\ wealth\ by\ top\ x\%\ firms \\ &if\ Wealth_i > 0 \end{aligned} \tag{17}$$

Where n=1 is the firm which created the most wealth and N is the company that created the least, but still positive, wealth. The proportion of gross wealth created by the top 2.5% of firms is used as the dependent variable when testing hypotheses 7, 8 and 9. To recap, these hypotheses infer that short-term return characteristics, corporate performance and its variability can explain differences in concentration of wealth creation.

### 5.5 Measuring concentration of net wealth creation across US industries

Concentration of net wealth creation is calculated differently from gross wealth concentration. This paper could calculate it similarly. However, because of significant wealth destruction by numerous companies in some industries the top 0.5% or top 1% of firms would already create

100% of net wealth. Additionally, proportions can go above 1 which would render using fractional outcome regressions impossible (why these are used is explained in the next section). Therefore, this paper measures what cumulative percentage of firms created 100% of the net wealth in their respective industries. For example, say there are 1000 firms in the retail industry and $100 billion of net wealth is created. After ranking all the firms from most wealth created to most wealth destructed and creating a cumulative sum of net wealth created, one could find that the top 30 firms are responsible for 100% of the net wealth created. In that case, 3% of firms are responsible for 100% of the net wealth created. To improve the intuitive understanding of this metric Table 12 shows how the top 3 firms or top 3% of firms create 100% of the net wealth in this sample. In this case the proportion of firms creating 100% of the net wealth is 0.03.

Table 12: Concentration of net wealth creation example

| PERMNO | Wealth created ($M) | Proportion of net wealth | Rank | Cumulative proportion of net wealth |
|---|---|---|---|---|
| 15 | 1060 | 0.53 | 1 | 0.53 |
| 37 | 620 | 0.31 | 2 | 0.84 |
| 86 | 320 | 0.16 | 3 | 1.00 |
| ... | ... | ... | ... | ... |
| 61 | 20 | 0.01 | 49 | 1.63 |
| 43 | 18 | 0.009 | 50 | 1.639 |
| 84 | 14 | 0.007 | 51 | 1.646 |
| ... | ... | ... | ... | ... |
| 5 | -100 | -0.05 | 98 | 1.18 |
| 98 | -160 | -0.08 | 99 | 1.10 |
| 77 | -200 | -0.10 | 100 | 1.00 |

Concentration of net wealth creation is captured in equation (18). Where the upper limit of the sum in the nominator, *n*100% NW*, is the marginal firm that makes the top group achieve 100% of net wealth creation. The outcome of the equation which is the proportion of firms responsible for 100% of the net wealth creation is a measure for wealth concentration across industries. These proportions are used when testing hypotheses 6, 7, 8 and 9. Hypothesis 6 states that a high degree of concentration in wealth creation, i.e. $< 5\%$ of stocks are responsible for 100% of net wealth creation, is present across US industries. A t-test is carried out to test the null-hypothesis that a proportion of 0.05 of stocks is responsible for all net wealth creation across US industries. Additionally, proportions describing concentration of net wealth creation are used as dependent variables when testing hypotheses 7, 8 and 9. Summed up, these hypotheses infer that short-term return characteristics, corporate performance and its variability can explain differences in wealth concentration.

$$\frac{\sum_{i=1}^{n_{100\%\ NW}} Wealth_i}{\sum_{i=1}^{N} Wealth_i} = Proportion\ of\ firms\ responsible\ for\ 100\%\ net\ wealth \tag{18}$$

## 5.6 Independent variables

This section outlines which independent variables are used to test hypotheses 3-5 and 7-9. First, hypotheses 3 and 7 use independent variables that resemble short-term return characteristics such as volatility, skewness and mean returns. Second, hypotheses 4 and 8 use independent variables that resemble industry corporate performance measures such as capitalization and asset turnover. Third, hypotheses 5 and 9 use independent variables that resemble variability of corporate performance within firms such as the average standard deviation of operating margins and interest coverage.

### 5.6.1 Short-term return characteristics

Compounding short-term returns over a long horizon can result in extreme returns. For example, a 0.5% monthly return that is compounded over 90 years results in a 21,745% return. Whereas a 0.5% simple return, that is 0.5% is earned every month over a fixed sum yields a 540% return over the same period. In this example, the effect of compounding results in a 40 times higher return.

However, in a real world setting such a constant short-term return over such a long horizon is highly unlikely, especially for stocks. Short-term returns tend to vary across stocks and over time. The degree of variation in returns is often called volatility in finance. Volatility can be calculated from the standard deviation of compound returns. Bessembinder (2018) provides evidence from simulation that compounding induces skewness in long-term returns, and that skewness depends on volatility. Farago and Hjalmarsson (2019) construct a theoretical framework confirming the finding that skewness in long-term returns is mainly driven by short-term volatility.

Skewness in long-term returns implies that a few firms have extreme returns. These firms are likely responsible for the majority of wealth creation. Thus, concentrated wealth creation might be caused by skewness in long-term returns, of which volatility is could be the main driver. Standard deviation is the underlying input for volatility, which is why standard deviation of short-term stock returns is one of the independent variables in the regressions of this paper. The data for the short-term stock return volatility variable is calculated with equation (19):

$$s_{industry} = \frac{\sum_{i=1}^{N} s_i}{N} \tag{19}$$

Where $s_{industry}$ is the average standard deviation of the industry, $s_i$ is the standard deviation of the monthly returns of a stock in that industry and $N$ is the number of stocks in that industry.

Another driving factor of long-term return skewness can be, quite intuitively, short-term return skewness. One can image that if short-term returns are skewed long-term returns will also be skewed. Short horizon positive skewness implies that very large positive short-term returns exist, which if compounded, could lead to skewness in long-term returns. Therefore, short-term skewness is also included in the regression analysis as an independent variable. This is implemented by calculating the average skewness of monthly returns of every stock in each industry as in equation (21):

$$g_{industry} = \frac{\sum_{i=1}^{N} g_i}{N} \tag{20}$$

Where $g_{industry}$ is the average monthly stock return skewness of a given industry, $g_i$ is the skewness of the monthly returns of a stock $i$ in that industry and $N$ is the number of stocks in that industry.

Besides standard deviation and skewness, this paper also includes mean monthly returns in the regression. Mean returns act as a control variable. The inputs for this variable are calculated by the following equation:

$$\bar{x}_{industry} = \frac{\sum_{i=1}^{N} \bar{x}_i}{N} \tag{21}$$

Where $\bar{x}_{industry}$ is the average mean monthly return of the industry, $\bar{x}_i$ is the mean monthly returns of a stock in that industry and $N$ is the number of stocks in that industry.

### 5.6.2 Corporate performance measures

Hypotheses 4 and 8 infer that corporate performance is an empirical determinant for the proportion of stocks that outperform the risk-free rate and concentration of wealth creation across US industries. Therefore, seven financial ratios are used as independent variables to test hypotheses 4 and 8. Section 4.5 details that capitalization ratio, asset turnover, quick ratio, return on capital employed, operating margin, cash flow/debt and interest coverage are used. This paper averages the median monthly financial ratios over the period from 1970 to 2018.

### 5.6.3 Variability in corporate performance

Hypotheses 5 and 9 infer that variability in corporate performance within firms has an effect on the degree of concentration and stocks' performance relative to the risk-free asset. Data to construct independent variables is obtained by using raw monthly firm financial ratios from WRDS. Ratios are winsorized at 1% to reduce outliers. To measure variability, standard deviations of financial ratios to each firm are calculated. Subsequently, standard deviations of firms within industries are averaged to obtain an industry level measurement of variability in corporate performance. Table 29 in Appendix B summarizes variability in corporate performance for every industry.

## 5.7 Fractional outcome regressions

To formally test hypotheses 3-5 and 7-9 fractional outcome regressions are used, because the dependent variables are proportions. Proportions have an interval from zero to one, i.e. [0,1]. Thus, the dependent variable is bounded. Consequently, as Papke and Wooldridge (1996) note, an ordinary least squares (OLS) regression cannot guarantee predictions to fall in the unit interval. Baum (2008) follows up on work by Papke and Wooldridge (1996) and provides a solution to this problem, which is to perform a logit transformation that makes sure the dependent variables are strictly within the unit interval:

$$y = \frac{1}{1 + exp(-X\beta)} \tag{22}$$

In equation (22), $y$ is the dependent variable, in this paper's case a proportion. For example the proportion of stocks outperforming the risk-free rate in a particular industry. $X$ is a set of independent variables and $\beta$ are the coefficients. The equation is equal to the more often used inverse logit transformation:

$$y = \frac{exp(X\beta)}{1 + exp(X\beta)} \tag{23}$$

Equations (22) and (23) can be rewritten into an equation where the dependent variable is transformed instead of the independent variables:

$$y* = log(\frac{y}{1-y}) = X\beta + \epsilon \tag{24}$$

Following Baum (2008), ε, a stochastic error process is added to equation (24). This allows the model to be fitted. Subsequently, a linear regression can be used to model *y**. After running those regressions one can let a statistical program such as Stata predict *y** and perform an inverse logit transformation on those values to get back to values on the interval (0,1). However, there is one major problem with this approach. Values of 0 and 1 cannot be fitted. For values of *y=0*, *log(0)* is not defined. Additionally, for values of *y=1*, we cannot divide by zero.

Several solutions to this problem have been proposed. The most rigorous solution is to drop all observations with 0 and 1, which is undesirable as one would lose observations. Another solution is to winsorize these observations with an arbitrary value of 0.001 or 0.999. However, Papke and Wooldridge (1996) proposed a more elegant solution that does not force us to alter inputs from the sample. The independent variables in the equation are transformed by a function $G(\cdot)$ to ensure that the predicted values of $y$ are in the interval (0,1).

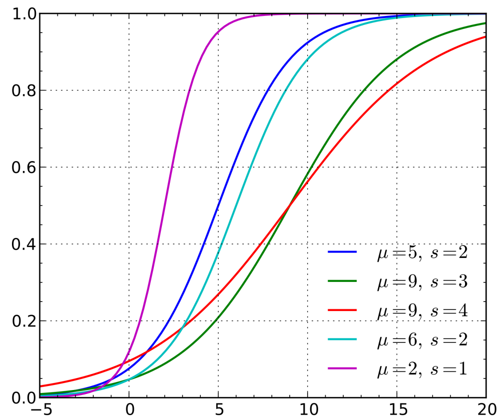
$$y = G(X\beta) \tag{25}$$

Where $G(\cdot)$ is a function that satisfies $0 < G(X\beta) < 1$ for all $X\beta \in \mathbb{R}$. There are several options of what function $G(\cdot)$ actually is. Logit transformation such as in equation (22) is a good candidate for $G(\cdot)$ as it ensures $y$ is always between 0 and 1. This type of function is a cumulative distribution function (CDF). Figure 1 displays several examples of logit CDFs. Note how the function is asymptotic to 0 and 1[6].

In order to estimate an equation such as equation (25), Papke and Wooldridge (1996) propose to follow earlier work from Gourieroux et al. (1984) and McCullagh and Nelder (1989). They outline a method called quasi-likelihood maximization. Going in-depth on all the econometric proofs of this method is outside the scope of this paper. Therefore, this paper aims to provide an intuitive understanding of the method used. To provide an intuitive understanding of this method, this paper breaks the method down into four pieces of statistics.

[6] Figure 1 is obtained from Wikipedia and created by user Krishnavedala.

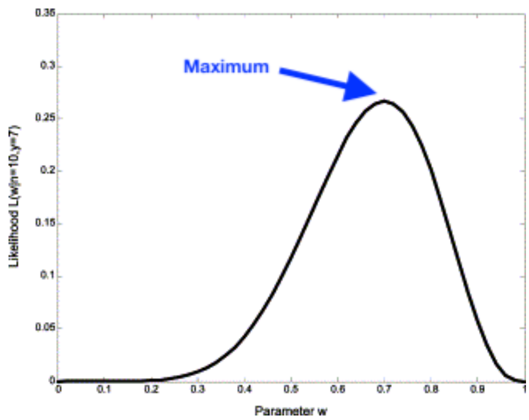
Figure 1: Cumulative distribution function logit




First, as described by Vail and Valavanis (1959), a likelihood function calculates how well a statistical model fits in relation to the sample data. Hence, as Myung (2003) points out, the likelihood function helps us in determining which version of a model is most likely to produce the data observed in the sample. Note that it does not give an answer as to which model in particular has the best fit. It is merely a tool that helps us in finding the best fit.

Second, finding the model with the best fit is the goal. In order to find that model, we need to maximize the likelihood function. To intuitively understand this, Figure 2 displays a very simple likelihood function[7]. The goal is to find the maximum indicated by the blue arrow. By taking the derivative of the likelihood function we can find the maximum.

Figure 2: Likelihood function example




Third, in many cases and also in this paper, the likelihood function is not as simple as the one displayed in Figure 2. Therefore, it can be quite hard to take the derivative of the likelihood function and find a solution for the maximum. To tackle this problem, *log* can be used. It is mathematically more convenient to maximize the log-likelihood function, which is simply the natural logarithm of the likelihood function. It would be quite elegant if we could take the derivative of the log-likelihood function. However, in most cases, as Myung (2003) points out, the model is too complex and non-linear. Therefore, a trial and error approach is more

[7]Figure 2 is borrowed from Myung (2003) tutorial on maximum likelihood estimation.

desirable. Computers are well equipped to perform such a task. A statistical program such as Stata, which is used by this paper, takes an iterative approach. The first iteration is either a guess or random. Each iteration takes into account the results from the previous iteration until it finds the best model. "The best model" is defined according to some criteria set by the person looking. For example, a criteria could be that a minimum or maximum number of iterations are performed. In our case Stata uses has a complex algorithm to determine what the best model is and generally uses 5 iterations.

The fourth and final step is to explain what *quasi* means in the context of quasi-likelihood maximization. Wedderburn (1974) states that it is often difficult to determine what the true distribution of the observations in a sample is. However, specifying the distribution is a requirement of maximizing likelihoods. Quasi-likelihoods solve this problem, because one does not have to specify the distribution but only needs to specify the form of the mean-variance relationship according to Wedderburn (1974). In this paper the form of the mean-variance relationship is a logit described in equation (23).

Having an intuitive understanding of how independent variables will be regressed on the dependent variables is paramount. To further support the understanding of the reader, equation (26) shows the general log-likelihood function that is maximized as described in StataCorp (2017).

$$ln(L) = \sum_{j=1}^{N} y_j ln\{G(x_j^{'}\beta)\} + (1 - y_j) ln\{1 - G(x_j^{'}\beta)\} \tag{26}$$

Where $N$ is the sample size, $y_j$ is the dependent variable, $ln(L)$ is maximized and $G(\cdot)$ is the functional form, logit, as described in equation (23). Which in this specific case is:

$$\frac{exp(x_j^{'}\beta)}{1 + exp(x_j^{'}\beta)} \tag{27}$$

Where $x_j$ are the covariates for individual $j$ and $\beta$ are the coefficients.

### 5.8 Regression equations

This section provides the regression equations that are used to test hypotheses 3-5 and hypotheses 7-9.

**Hypothesis 3:** short-term stock return volatility, stock return skewness and mean stock returns can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset. Equation (28) is used to test hypothesis 3.

$$E(proportion\ R > Rf|x) = G(\beta_0 + \beta_1 s_{industry} + \beta_2 g_{industry} + \beta_3 \bar{x}_{industry}) \tag{28}$$

The dependent variable $proportion\ R > Rf$ is the proportion of firms that outperform the risk-free rate over their lifetime in a given industry. $G(\cdot)$ is the logistic transformation to ensure estimated values for $E(proportion\ R > Rf|x)$ are in the range [0,1]. $\beta_0$ is the constant. $s_{industry}$

is the average standard deviation of monthly stock returns in a given industry. $g_{industry}$ is the average skewness of monthly stocks returns in a given industry. $\bar{x}_{industry}$ is the average mean monthly return of a given industry. Based on the arguments provided in the hypothesis development section 3 we expect the coefficients for $s_{industry}$ and $g_{industry}$ to have a negative sign. A coefficient with positive sign is expected for $\bar{x}_{industry}$.

**Hypothesis 4:** corporate performance can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset. Equation (29) is used to test hypothesis 4.

$$\begin{aligned} &E(proportion\ R > Rf|x) = G(\beta_0 + \beta_1 Capital\ ratio \\ &+ \beta_2 Asset\ turnover + \beta_3 Quick\ ratio \\ &+ \beta_4 ROCE + \beta_5 Operating\ margin \\ &+ \beta_6 Cash\ flow/debt + \beta_7 Interest\ coverage) \end{aligned} \tag{29}$$

The dependent variable and logistics transformation are the same as in equation (28). How the seven independent variables that measure corporate performance are calculated is specified in section 5.6.2. Based on the arguments presented in the hypothesis development section 3 we expect the coefficient of $Capital\ ratio$ to have a negative sign and coefficients of all the other independent variables to have a positive sign.

**Hypothesis 5:** variability in corporate performance can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset. Equation (30) is used to test hypothesis 5.

$$\begin{aligned} &E(proportion\ R > Rf|x) = G(\beta_0 + \beta_1 Capital\ ratio\ SD \\ &+ \beta_2 Asset\ turnover\ SD + \beta_3 Quick\ ratio\ SD \\ &+ \beta_4 ROCE\ SD + \beta_5 Operating\ margin\ SD \\ &+ \beta_6 Cash\ flow/debt\ SD + \beta_7 Interest\ coverage\ SD) \end{aligned} \tag{30}$$

The dependent variable and logistics transformation are the same as in (28). How the seven independent variables that measure variability corporate performance are calculated is specified in the earlier section 5.6.3. Based on the arguments presented in the hypothesis development section 3 we expect the coefficients of all independent variables to have a negative sign.

**Hypothesis 7:** short-term stock return volatility, return skewness and mean returns can explain differences in the proportion of stocks that is responsible for a $x\%$ of wealth across US industries. Equation (31) is used to test hypothesis 7.

$$E(proportion\ wealth\ concentration|x) = G(\beta_0 + \beta_1 s_{industry} + \beta_2 g_{industry} + \beta_3 \bar{x}_{industry}) \tag{31}$$

The dependent variable $E(proportion\ wealth\ concentration|x)$ measures the degree of concentration in either gross or net wealth creation. Separate regressions are executed for concentration in gross and net wealth creation.

Concentration in gross wealth creation is measured by what proportion of wealth is created by the top 2.5% of firms. Therefore, when a coefficient has a **positive** sign it means that a variable is associated with more concentration in gross wealth creation. Hence, regarding the regressions of concentration of gross wealth creation we expect coefficients for $s_{industry}$ (return volatility) and $g_{industry}$ (return skewness) to have a positive sign whereas we expect the coefficient for $\bar{x}_{industry}$ (mean return) to have a negative sign.

Concentration in net wealth creation is measured by what proportion of stocks is responsible for all net wealth creation. Therefore, when a coefficient has **negative** sign it means that a variable is associated with more concentration in net wealth creation. Hence, regarding the regressions of concentration in net wealth creation we expect coefficients for $s_{industry}$ (return volatility) and $g_{industry}$ (return skewness) to have a negative sign whereas we expect the coefficient for $\bar{x}_{industry}$ (mean return) to have a positive sign.

**Hypothesis 8:** corporate performance can explain differences in the proportion of stocks that is responsible for a $x$% of wealth across US industries. Equation (32) is used to test hypothesis 8.

$$\begin{aligned} &E(proportion\ wealth\ concentration|x) = G(\beta_0 + \beta_1 Capital\ ratio \\ &+ \beta_2 Asset\ turnover + \beta_3 Quick\ ratio + \beta_4 ROCE + \beta_5 Operating\ margin \\ &+ \beta_6 Cashflow/debt + \beta_7 Interest\ coverage) \end{aligned} \tag{32}$$

The dependent variable $E(proportion\ wealth\ concentration|x)$ measures the degree of concentration in either gross or net wealth creation. Separate regressions are executed for concentration in gross and net wealth creation.

Concentration in gross wealth creation is measured by what proportion of wealth is created by the top 2.5% of firms. Therefore, when a coefficient has a **positive** sign it means that a variable is associated with more concentration in gross wealth creation. Hence, regarding the regression of concentration in gross wealth creation we expect the coefficient for $Capital\ ratio$ to have a positive sign and coefficients of all other independent variables to have a negative sign. In other words, better median corporate performance of firms in a given industry is associated with less concentration of wealth creation.

Concentration in net wealth creation is measured by what proportion of stocks is responsible for all net wealth creation. Therefore, when a coefficient has **negative** sign it means that a variable is associated with more concentration in net wealth creation. Hence, regarding the regression of concentration in net wealth creation we expect the coefficient of $Capital\ ratio$ to have a negative sign and coefficients of all other independent variables to have a positive sign. In other words, better median corporate performance of firms in a given industry is associated with less concentration of wealth creation.

**Hypothesis 9:** variability in corporate performance can explain differences in the proportion of stocks that is responsible for a $x$% of wealth across US industries. Equation (33) is used to test hypothesis 9.

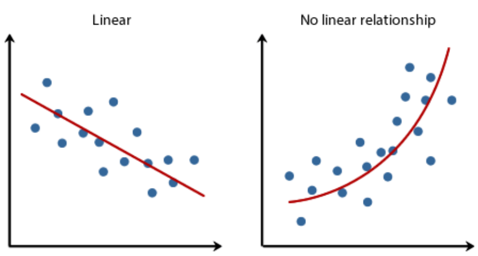
$$E(proportion\ wealth\ concentration|x) = G(\beta_0 + \beta_1 Capital\ ratio\ SD$$
$$+ \beta_2 Asset\ turnover\ SD + \beta_3 Quick\ ratio\ SD + \beta_4 ROCE\ SD \tag{33}$$
$$+ \beta_5 Operating\ margin\ SD + \beta_6 Cash\ flow/debt SD + \beta_7 Interest\ coverage\ SD)$$

The dependent variable $E(proportion\ wealth\ concentration|x)$ measures the degree of concentration in either gross or net wealth creation. Separate regressions are executed for concentration in gross and net wealth creation. Regarding the regression of concentration in gross wealth creation we expect all variables to have a coefficient with a negative sign. In other words, more variability in the corporate performance of firms in a given industry is associated with less concentration in gross wealth creation. Regarding the regression of concentration in net wealth creation we expect all variables to have a coefficient with a positive sign. In other words, more variability in the corporate performance of firms in a given industry is associated with less concentration in net wealth creation.

### 5.9 Interpreting fractional regression outputs: marginal effects

This section explains how coefficients from fractional outcome regressions can be interpreted in a practical and meaningful manner. In a linear regression, the coefficient of an independent variable is used to measure an effect on the dependent variable. For example, if the coefficient of *operating margin* is $\beta_2 = 0.05$ we know that for a one unit increase in *operating margin*, all else equal, the conditional expected value of $y$ increases by 0.05. This is what Cameron and Trivedi (2010) describe as the ”marginal effect” of an independent variable. However, non-linear models, such as the logistic regression used in this paper, cannot be plainly interpreted from the regression output, since they have undergone a transformation. Figure 3 shows that a linear relationship has a constant slope, whereas a non-linear relationship has a varying slope[8]. To intuitively understand this, a unit increase in *operating margin* with coefficient $\beta_2 = 0.05$ does not lead to a 0.05 increase of the conditional expected value $y$. Instead, it leads to an increase that is logistically transformed. More importantly, a one unit change from 1 to 2 yields a different change in the expected value of $y$ than a change from 9 to 10, because the equation in non-linear.

Figure 3: Linear vs. nonlinear relationship




To solve this problem, one could calculate the marginal effect for a range of values of *operating margin*, while keeping the other independent variables constant at some chosen values. In essence, what

[8]Figure 3 is kindly borrowed from Laerd Statistics.

one would do is calculate the rate of change (or the slope if you will) for a range of values of *operating margin*. Since the equation is non-linear, one would observe that the rate of change at $operating\ margin = 1$ is different than at $operating\ margin = 9$. To make the results more insightful one could plot all the marginal effects measured. However, there is a flaw with this approach, because for which case or under what circumstances is one actually calculating the marginal effects of the independent variable? Some values for the other independent variables must be chosen. An option is to keep other independent variables constant when calculating the marginal effects for *operating margin*. For example, by assuming that other independent variables such as ROCE[9] is constant at $ROCE = 10$. However, is $ROCE = 10$ a typical case or an outlier? These are important questions that need to be answered to make the marginal effects resemble what is seen in practice.

Williams (2012) explains how this problem can be solved. The most simple solution is to let other independent variables take the mean observed in the sample. Essentially, creating an "average" case. For instance, mean $ROCE = 5$ and mean $Standard\ deviation = 0.20$ across industries. Then, for values in a range of *operating margin* one would calculate the marginal effect while keeping the other variables at their mean. The result is that one would know how the expected value of the dependent variable is influenced by *operating margin* for the average industry. This approach is called marginal effects at the means (MEMs). However, the approach is not ideal, because we would only know how the average industry is influenced by a certain independent variable. Additionally, the mean values for the other independent variables might not even exist in practice as Williams (2012) points out. Hence, the marginal effects at the means approach is not ideal.

There must be a better way. Royston (2013) and Williams (2012) highlight an approach where all the underlying observations are used from the sample called average marginal effects (AMEs). Imagine we only want to calculate the marginal effect of the independent variable *operating margin* at a fixed value say $operating\ margin = 15$. In the previous approach one would take the mean for the other independent variables and see what the marginal effect is at $operating\ margin = 15$. However, in AMEs one would go to the first case in the sample and pretend $operating\ margin = 15$, use the observed values from the first case as the input for the other variables and calculate the marginal effect. One would then go to the second case, again pretend $operating\ margin = 15$, use the observed values from the second case as the input for the other variables and calculate the marginal effect. This process is repeated for every case in the sample. The result is a list of marginal effects at $operating\ margin = 15$ calculated with all the observed values for the other variables in the sample. The next step is to take the average of all the marginal effects calculated, or in other words the average marginal effect (AME). Thus, the average marginal effect is more representative for actual observation made in the real world.

Finally, the AMEs have to be calculated for a whole range of values of the independent variable one is interested in. For example, we not only want to know what the marginal effect is of $operating\ margin = 15$, but also for a whole range of other values of *operating margin*.

[9] As specified in previous sections, ROCE stands for return on capital employed and is a measure of profitability.

Therefore, a range is specified for the independent variable over which we are interested to know the marginal effects. For example, from $operating\ margin = -10$ until $operating\ margin = 40$. Hence, to intuitively understand reported coefficients from regression analysis we plot marginal effects.

# 6 Empirical results

## 6.1 Stocks' ability to outperform the risk-free asset

Hypothesis 1 has been tested in section 4.3. Therefore, this section opens with reporting the results of hypothesis 2. Hypothesis 2 states that the majority of stocks across US industries are unable to achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset. Table 13 reports the percentage of stocks in each industry with a higher buy-and-hold return over their lifetime relative compared to several benchmarks. Stocks' performance against benchmarks varies significantly across industries. As can be seen from the table, just 42% of stocks outperform the matched return of the Treasury bill in the period from 1926 to 2018. Additionally, just 44% of stocks outperform inflation (CPI) over their lifetime. Large differences occur across industries, for example, 74% of stocks in the utilities industry manage to beat the Treasury bill over their lifetime. On the contrary, just 22% of stocks in the precious metal industry outperform the Treasury bill over their lifetime. In 15 out of 50 industries, more than 50% of stocks manage to have a higher buy-and-hold return over their lifetime than the risk-free rate. Hence, from visual inspection of Table 13, it is likely that we fail to reject hypothesis 2. Hypothesis 2 is tested in a single sample t-test to determine whether the proportion of stocks that have a lifetime buy-and-hold return exceeding the return of the risk-free asset was lower than 0.5 across US industries. Mean proportion of stocks outperforming the risk-free asset over their lifetime ($0.44 \pm 0.12$) was significantly lower than 0.5, t(49)=-3.1706, p=0.0013[10]. Thus, we fail to reject hypothesis 2.

The degree of underperformance increases as horizons become longer. Table 14 reports the benchmarked performance of stocks across industries on a monthly basis. Although differences are already visible between industries, they are less pronounced over this shorter horizon. Additionally, Table 14 shows that even over such a short horizon stocks tend to underperform Treasury bills and inflation. This could be caused by short-term volatility which is higher for stocks than for Treasury bills and the rate of inflation. Hence, the probability of a negative return is larger for stocks than for the benchmarks.

Performance relative to benchmarks seems to be getting worse in more recent decades. Table 35 in Appendix C displays outperformance figures of individual stocks across industries from 1970 until 2018. It is apparent that an even lower percentage of stocks manage to beat the benchmarks. For example, just 19% of stocks in the recreation industry beat the value-weighted market return over their lifetime. Section 9 in Appendix C contains additional tables on outperformance for monthly, annual, decade and lifetime horizons for both the period 1926 to 2018 and 1970 to 2018.

[10] p-value evaluates whether the proportion was lower than 0.5

Evidence for hypothesis 1, which was tested in section 4.3, is reinforced by analyzing performance of individual stock returns against market and industry returns. Table 13 reports what percentage of stocks in each industry is able to exceed the matched market and industry return over its lifetime. Regarding market returns, a majority of stocks in all but two industries (candy & soda and banking) do not outperform the equal- or value-weighted market return. When benchmarking individual stock returns against industry returns this paper observes a similar pattern. In the vast majority of industries stocks have a harder time beating the equal-weighted benchmarks as opposed to the value-weighted benchmarks. In value-weighted market portfolios, smaller firms have a lower weight than larger firms since their market capitalization is lower. Therefore, it is likely that the size effect, where small firm outperform large firms, as described by Fama and French (1993), is also present in our sample.

How can the long-term buy-and-hold return of the market be higher than the Treasury bill and inflation if the majority of stocks have a lower return than the Treasury bill and inflation in both the short- and long-term? The simple explanation is that a minority of stocks perform so well that they significantly increase the mean return of the overall market. This section emphasises the performance relative to the Treasury bill and inflation, because having excess return over Treasury bills or inflation is a requirement for wealth creation. This paper found very high skewness in stock returns and has shown that the majority of stocks underperform Treasury bills and inflation. Therefore, it is likely that wealth creation is concentrated in a number of industries, which is tested in subsequent sections.

Table 13: Percentage of stocks that outperform their benchmark over a lifetime horizon during 1926-2018

| | **Non-stock benchmarks** | | | **Market benchmarks** | | **Industry benchmarks** | |
|---|---|---|---|---|---|---|---|
| | % > 0 | % > T-bill | % > CPI | % > EW Mkt return | % > VW Mkt return | % > EW Ind return | % > VW Ind return |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 40 | 33 | 34 | 17 | 25 | 27 | 24 |
| Food Products | 64 | 57 | 58 | 33 | 40 | 29 | 39 |
| Candy & Soda | 72 | 65 | 64 | 41 | 51 | 31 | 49 |
| Beer & Liquor | 61 | 50 | 53 | 30 | 31 | 33 | 27 |
| Tobacco Products | 69 | 65 | 67 | 31 | 44 | 21 | 40 |
| Recreation | 35 | 27 | 29 | 17 | 20 | 26 | 24 |
| Entertainment | 36 | 32 | 33 | 21 | 25 | 25 | 23 |
| Printing & Publishing | 56 | 47 | 49 | 31 | 38 | 28 | 36 |
| Consumer Goods | 51 | 39 | 42 | 24 | 30 | 27 | 32 |
| Apparel | 47 | 39 | 40 | 24 | 30 | 31 | 32 |
| Healthcare | 45 | 38 | 40 | 29 | 30 | 29 | 33 |
| Medical Equipment | 44 | 38 | 39 | 26 | 27 | 27 | 23 |
| Pharmaceutical Products | 38 | 35 | 35 | 27 | 27 | 24 | 24 |
| Chemicals | 61 | 55 | 56 | 33 | 41 | 31 | 42 |
| Rubber and Plastic Products | 54 | 45 | 45 | 30 | 33 | 30 | 39 |
| Textiles | 54 | 45 | 45 | 22 | 30 | 32 | 34 |
| Construction Materials | 62 | 53 | 55 | 29 | 39 | 28 | 41 |
| Construction | 44 | 35 | 38 | 22 | 26 | 29 | 30 |
| Steel Works | 60 | 51 | 52 | 22 | 30 | 28 | 40 |
| Fabricated Products | 54 | 39 | 44 | 25 | 32 | 30 | 39 |
| Machinery | 57 | 49 | 52 | 30 | 35 | 29 | 41 |
| Electrical Equipment | 40 | 33 | 35 | 21 | 23 | 23 | 24 |
| Automobiles and Trucks | 53 | 45 | 46 | 21 | 27 | 28 | 29 |
| Aircraft | 66 | 59 | 59 | 36 | 45 | 26 | 35 |
| Shipbuilding, Railroad Equipment | 58 | 52 | 52 | 28 | 35 | 30 | 36 |
| Defense | 60 | 53 | 55 | 36 | 45 | 34 | 32 |
| Precious Metals | 27 | 22 | 24 | 16 | 14 | 20 | 23 |
| Mines | 40 | 33 | 35 | 24 | 26 | 29 | 31 |
| Coal | 57 | 49 | 51 | 25 | 31 | 35 | 39 |
| Petroleum and Natural Gas | 38 | 32 | 34 | 19 | 24 | 24 | 23 |
| Utilities | 80 | 74 | 75 | 38 | 48 | 35 | 56 |
| Communication | 46 | 41 | 42 | 31 | 32 | 31 | 33 |
| Personal Services | 43 | 36 | 36 | 26 | 28 | 30 | 34 |
| Business Services | 45 | 38 | 40 | 29 | 30 | 28 | 31 |
| Computers | 35 | 27 | 29 | 18 | 20 | 21 | 19 |
| Computer Software | 35 | 32 | 33 | 24 | 25 | 23 | 22 |
| Electronic Equipment | 43 | 36 | 37 | 24 | 26 | 21 | 30 |
| Measuring and Control Equipment | 51 | 40 | 43 | 25 | 29 | 21 | 28 |
| Business Supplies | 67 | 61 | 63 | 31 | 45 | 35 | 49 |
| Shipping Containers | 74 | 64 | 65 | 41 | 48 | 34 | 48 |
| Transportation | 49 | 44 | 44 | 26 | 30 | 32 | 34 |
| Wholesale | 44 | 38 | 39 | 28 | 29 | 29 | 29 |
| Retail | 48 | 40 | 42 | 26 | 32 | 30 | 32 |
| Restaurant, Hotels, Motels | 45 | 38 | 39 | 25 | 27 | 30 | 29 |
| Banking | 70 | 65 | 66 | 47 | 53 | 44 | 54 |
| Insurance | 67 | 57 | 58 | 38 | 44 | 35 | 44 |
| Real Estate | 55 | 49 | 50 | 34 | 33 | 37 | 37 |
| Trading | 70 | 62 | 64 | 47 | 45 | 39 | 41 |
| Other | 42 | 38 | 37 | 29 | 26 | 32 | 34 |
| Not available | 46 | 42 | 40 | 31 | 30 | 39 | 36 |
| Total | 51 | 42 | 44 | 27 | 30 | 29 | 33 |

All lifetime buy-and-hold common stock and REIT stock returns from January 1926 until December 2018 are benchmarked against zero, 30-day Treasury Bills, the Consumer Price Index (CPI), equal- and value-weighted market returns and equal- and value-weighted industry returns. Equal-weighted (EW) means that all stocks are given the same weight when calculating market and industry returns. Value-weighted (VW) means that stocks are given a weight based on their market capitalization when calculating market and industry returns. Stocks are allowed to switch industries multiple times. If stocks switch back to the industry they were previously in, this is treated as a separate run of returns to prevent faulty linking of returns. Buy-and-hold stock returns are exactly matched to the benchmark on a monthly frequency.

Table 14: Percentage of stocks that outperform their benchmark over a monthly horizon during 1926-2018

| | **Non-stock benchmarks** | | | **Market benchmarks** | | **Industry benchmarks** | |
|---|---|---|---|---|---|---|---|
| | % > 0 | % > T-bill | % > CPI | % > EW Mkt return | % > VW Mkt return | % > EW Ind return | % > VW Ind return |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 43 | 42 | 43 | 43 | 44 | 46 | 45 |
| Food Products | 50 | 49 | 49 | 46 | 47 | 46 | 47 |
| Candy & Soda | 52 | 51 | 51 | 48 | 48 | 47 | 48 |
| Beer & Liquor | 48 | 47 | 48 | 46 | 46 | 47 | 46 |
| Tobacco Products | 51 | 50 | 50 | 47 | 48 | 48 | 48 |
| Recreation | 43 | 43 | 43 | 44 | 44 | 46 | 46 |
| Entertainment | 44 | 44 | 44 | 45 | 45 | 46 | 45 |
| Printing & Publishing | 49 | 48 | 48 | 47 | 47 | 47 | 48 |
| Consumer Goods | 47 | 46 | 47 | 46 | 46 | 46 | 46 |
| Apparel | 46 | 46 | 46 | 45 | 45 | 47 | 46 |
| Healthcare | 46 | 46 | 46 | 46 | 46 | 46 | 47 |
| Medical Equipment | 45 | 45 | 46 | 45 | 45 | 46 | 45 |
| Pharmaceutical Products | 46 | 46 | 46 | 45 | 45 | 45 | 45 |
| Chemicals | 50 | 49 | 49 | 47 | 47 | 47 | 47 |
| Rubber and Plastic Products | 46 | 46 | 46 | 45 | 46 | 46 | 47 |
| Textiles | 46 | 46 | 46 | 45 | 46 | 46 | 47 |
| Construction Materials | 47 | 47 | 47 | 46 | 46 | 46 | 47 |
| Construction | 45 | 45 | 45 | 45 | 45 | 46 | 47 |
| Steel Works | 49 | 48 | 49 | 46 | 46 | 46 | 47 |
| Fabricated Products | 45 | 44 | 45 | 45 | 45 | 46 | 47 |
| Machinery | 49 | 48 | 48 | 46 | 47 | 46 | 47 |
| Electrical Equipment | 45 | 45 | 45 | 45 | 45 | 45 | 46 |
| Automobiles and Trucks | 48 | 47 | 48 | 45 | 46 | 46 | 47 |
| Aircraft | 49 | 48 | 49 | 47 | 47 | 46 | 47 |
| Shipbuilding, Railroad Equipment | 50 | 49 | 50 | 46 | 47 | 47 | 47 |
| Defense | 50 | 50 | 50 | 47 | 48 | 48 | 48 |
| Precious Metals | 40 | 40 | 41 | 42 | 42 | 46 | 45 |
| Mines | 43 | 43 | 44 | 44 | 44 | 46 | 46 |
| Coal | 47 | 46 | 47 | 45 | 46 | 47 | 47 |
| Petroleum and Natural Gas | 44 | 44 | 44 | 44 | 45 | 45 | 45 |
| Utilities | 55 | 53 | 54 | 48 | 49 | 48 | 50 |
| Communication | 48 | 47 | 48 | 46 | 46 | 46 | 46 |
| Personal Services | 46 | 45 | 46 | 45 | 46 | 47 | 47 |
| Business Services | 46 | 46 | 46 | 46 | 46 | 46 | 46 |
| Computers | 44 | 44 | 45 | 45 | 45 | 46 | 46 |
| Computer Software | 46 | 46 | 46 | 45 | 45 | 45 | 45 |
| Electronic Equipment | 46 | 45 | 46 | 45 | 46 | 46 | 46 |
| Measuring and Control Equipment | 45 | 45 | 46 | 45 | 46 | 45 | 46 |
| Business Supplies | 50 | 49 | 49 | 47 | 47 | 47 | 48 |
| Shipping Containers | 50 | 50 | 50 | 47 | 48 | 47 | 48 |
| Transportation | 48 | 47 | 48 | 45 | 46 | 47 | 47 |
| Wholesale | 46 | 45 | 46 | 45 | 45 | 46 | 46 |
| Retail | 48 | 47 | 47 | 46 | 47 | 47 | 47 |
| Restaurant, Hotels, Motels | 46 | 45 | 46 | 45 | 45 | 46 | 46 |
| Banking | 51 | 50 | 51 | 48 | 48 | 47 | 48 |
| Insurance | 51 | 50 | 50 | 48 | 48 | 47 | 48 |
| Real Estate | 49 | 48 | 49 | 47 | 47 | 48 | 47 |
| Trading | 50 | 49 | 49 | 47 | 47 | 46 | 46 |
| Other | 46 | 46 | 46 | 45 | 45 | 46 | 46 |
| Not available | 45 | 45 | 45 | 44 | 43 | 47 | 46 |
| Total | 46 | 46 | 46 | 45 | 46 | 46 | 47 |

All monthly buy-and-hold common stock and REIT stock returns from January 1926 until December 2018 are benchmarked against zero, 30-day Treasury Bills, the Consumer Price Index (CPI), equal- and value-weighted market returns and equal- and value-weighted industry returns. Equal-weighted (EW) means that all stocks are given the same weight when calculating market and industry returns. Value-weighted (VW) means that stocks are given a weight based on their market capitalization when calculating market and industry returns. Stocks are allowed to switch industries multiple times. If stocks switch back to the industry they were previously in, this is treated as a seperate run of returns to prevent faulty linking of returns. Buy-and-hold stock returns are exactly matched to the benchmark on a monthly frequency.

## 6.2 Evaluating cross-industry variation in performance against US T-bills

Hypotheses 3, 4 and 5 address the research question of this paper indirectly by asking what explains differences in proportions of stocks across US industries that are able to generate a higher return than the risk-free asset over their lifetime and are thus likely to create wealth. Answers to that question help solve the research question: "Do US industries suffer from concentration in wealth creation and what are the empirical determinants of cross-industry variation in concentration of wealth creation?" Table 15 reports outputs of fractional regressions used to test hypotheses 3, 4 and 5. Regressions (1) and (2) test hypothesis 3, regression (3) tests hypothesis 4 and regression (4) tests hypothesis 5. Regression (5) tests the hypotheses combined. The subsequent sections explain the results and interpret their economic meaning.

Hypothesis 3 states that short-term return volatility, return skewness and mean returns can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the risk-free asset. Regressions (1) and (2) from Table 15 test hypothesis 3 and find support for it. Coefficients for average standard deviation (volatility) are negative and significant at the 0.1% level. Hence, a higher average short-term return volatility in a given industry is associated with less stocks outperforming US Treasury bills over their lifetime (t-stat=-7.79; t-stat=-5.96). Figure 4 displays the marginal effects plot (hereinafter MEP) of volatility in the period 1970 to 2018[11]. This is in line with our expectations and with Bessembinder et al. (2019) who also find that short-term return volatility is associated with less stocks outperforming the matched risk-free rate over their lifetime.

From visual inspection of our sample we learn that "innovative" industries such as pharmaceutical, software and computer experience high stock return volatility. Gharbi et al. (2014) research the relation between stock return volatility and R&D intensity of high-tech firms in France and find a positive relation. In addition, Mazzucato and Tancioni (2012) use patents as a measure of innovation while investigating the stock-return volatility in the pharmaceutical industry. They find that innovation is positive associated with stock return volatility. Therefore, industries where firms are highly dependent on the success of their innovations could be more risky and experience higher short-term stock return volatility which leads to a lower proportion of firms that outperform the risk-free rate over their lifetime.

No significant explanatory power is found for the average skewness of short-term individual stock returns (t-stat=0.35; t-stat=-1.36). Farago and Hjalmarsson (2019) show through simulation that short-term skewness amplifies long-term skewness if volatility if sufficiently high. Hence, volatility in our sample might be too low for short-term skewness to increase long-term skewness to a point where a large proportion of stocks underperform the risk-free rate. Coefficients for mean returns are positive and significant at the 1% level. Thus, a higher average mean monthly return in a given industry is associated with a higher proportion of stocks outperforming US Treasury bills over their lifetime (t-stat=3.09; t-stat=3.73). Figure 5 plots the marginal effects of mean returns on the proportion of stocks that outperform US Treasury bills over their lifetime in the period 1970 to 2018. We expected this results and are of the opinion it is intuitive since

[11] The shaded area represents the 95% confidence interval.

Figure 4: MEP volatility

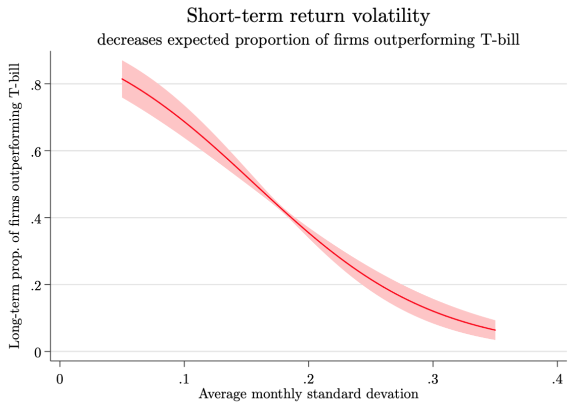
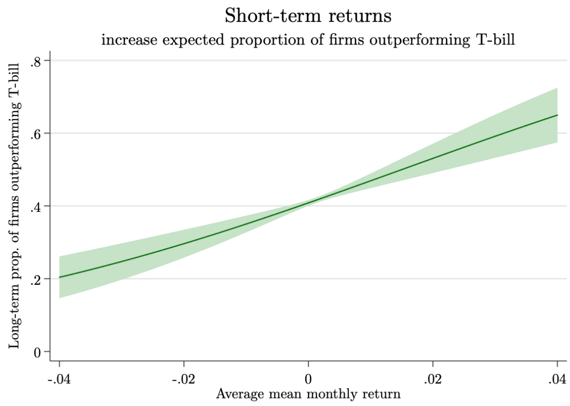



Figure 5: MEP mean return




beating a return benchmark and higher returns are directly related. In conclusion, we find strong evidence in support of hypothesis 3 and accept it. This conclusion remains the same when hypothesis 3 is tested alongside variables from other hypotheses in regression (5) .

Hypothesis 4 states that corporate performance can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset. Regression (3) from Table 15 tests hypothesis 4 and finds limited support for it. We expect better corporate performance in a given industry to be associated with a larger proportion of firms outperforming the risk-free rate over their lifetime. A higher average capital ratio, i.e. more leverage, in a given industry is associated with more stocks outperforming US Treasury bills over their lifetime (t-stat=2.34). This is contrary to what we expected in the hypothesis development. We expected industries with more leveraged firms to be prone to bankruptcies which could decrease the proportion of firms outperforming the risk-free rate. Visual inspection of the data learns that firms in the utilities and banking industry tend to carry substantial leverage and perform above-average compared to other industries with respect to outperforming the risk-free rate. Both industries are characterised by stringent regulation as default could damage consumers by wiping out savings or disrupting utility services. **?** provide a possible explanation by showing that regulation could shield banks from competition which might make them willing to take on more leverage. Hence, industry specific characteristics could be responsible for positive relation between capital ratio and the proportion of stocks that outperform the risk-free rate over their lifetime.

A higher average cash flow/debt in a given industry is associated with less stock outperforming US Treasury bills over their lifetime (t-stat=-3.32). This result might also be driven by an advantage in leverage or specific characteristics of high leverage industry. Other variables such as average asset turnover (t-stat=-1.38), quick ratio (t-stat=0.54), return on capital employed (t-stat=0.87), operating margin (t-stat=0.60) and interest coverage (t-stat=1.40) do not have significant explanatory power.

Regression (5) tests hypothesis 4 in combination with variables from other hypotheses 3 and 5. We find supportive evidence for hypothesis 4. Once again, a higher average capital ratio in a

Table 15: Cross-industry outperformance of T-bills

| | Proportion of firms outperforming T-bill over their lifetime | | | | |
|---|---|---|---|---|---|
| | 1926-2018 | 1970-2018 | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Standard deviation | $-11.65^{***}$ | $-9.408^{***}$ | | | $-14.02^{***}$ |
| | $(-7.79)$ | $(-5.96)$ | | | $(-9.54)$ |
| Skewness | 0.113 | $-0.435$ | | | 0.139 |
| | (0.35) | $(-1.36)$ | | | (0.87) |
| Mean return | $22.04^{**}$ | $21.75^{***}$ | | | $25.56^{***}$ |
| | (3.09) | (3.73) | | | (5.80) |
| Capital ratio | | | $2.109^{*}$ | | $1.109^{*}$ |
| | | | (2.34) | | (2.39) |
| Asset turnover | | | $-0.253$ | | $0.307^{**}$ |
| | | | $(-1.38)$ | | (2.76) |
| Quick ratio | | | 0.141 | | 0.112 |
| | | | (0.54) | | (1.08) |
| ROCE | | | 3.391 | | 0.338 |
| | | | (0.87) | | (0.26) |
| Operating margin | | | 0.593 | | $1.623^{**}$ |
| | | | (0.60) | | (3.05) |
| Cash flow/debt | | | $-2.935^{***}$ | | 0.558 |
| | | | $(-3.32)$ | | (1.12) |
| Interest coverage | | | 0.0779 | | $-0.0110$ |
| | | | (1.40) | | $(-0.48)$ |
| Capital ratio SD | | | | 1.564 | $-0.764$ |
| | | | | (0.39) | $(-0.40)$ |
| Asset turnover SD | | | | $-1.887$ | $-1.332$ |
| | | | | $(-1.40)$ | $(-1.71)$ |
| Quick ratio SD | | | | $-0.364$ | $-0.0188$ |
| | | | | $(-1.25)$ | $(-0.10)$ |
| ROCE SD | | | | $-4.561$ | $6.773^{***}$ |
| | | | | $(-1.48)$ | (4.60) |
| Operating margin SD | | | | $0.793^{*}$ | $0.501^{**}$ |
| | | | | (1.97) | (2.97) |
| Cash flow/debt SD | | | | $-4.292^{***}$ | $-2.102^{***}$ |
| | | | | $(-3.43)$ | $(-3.96)$ |
| Interest coverage SD | | | | $0.0230^{**}$ | 0.00172 |
| | | | | (2.83) | (0.46) |
| Constant | $1.688^{***}$ | $1.700^{***}$ | $-1.202$ | 0.978 | $0.877^{*}$ |
| | (6.78) | (8.23) | $(-1.62)$ | (1.92) | (2.55) |
| Observations | 50 | 50 | 49 | 50 | 49 |

$t$ statistics in parentheses

This fractional logistic regression measures the effect of the short-term mean, standard deviation (SD) and skewsness of stock returns on long-term outperformance over T-bills of individual stocks across industries. The mean, SD and skewness are calculated for each stock and averaged to get to the industry figures. In regressions (3), (4) and (5) corporate performance figures and the standard deviation of those figures is also taken into account.

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

given industry is associated with more stocks outperforming US T-bills (t-stat=2.39). Higher average efficiency (t-stat=2.76) and profitability (t-stat=3.05) in a given industry is also associated with a higher proportion of stocks outperforming US Treasury bills in the long run which is in line with our expectations. Haugen et al. (1996) find that higher efficiency in the form of asset turnover could indicate growth potential of a firm which could increase its chance of outperforming the risk-free rate. In addition, Fama and French (2015) recognize that more profitable firms tend to post higher returns than less profitable firms. Hence, there are economic arguments for the result that industries with more efficient and profitable firms tend to have a larger proportion of firms that outperform the risk-free rate over their lifetime. Figure 6 and 7 display marginal effects of both efficiency and profitability in a given industry on proportions of stocks that outperform US T-bills over their lifetime. In conclusion, this paper finds evidence in support of hypothesis 4.

Figure 6: MEP asset turnover

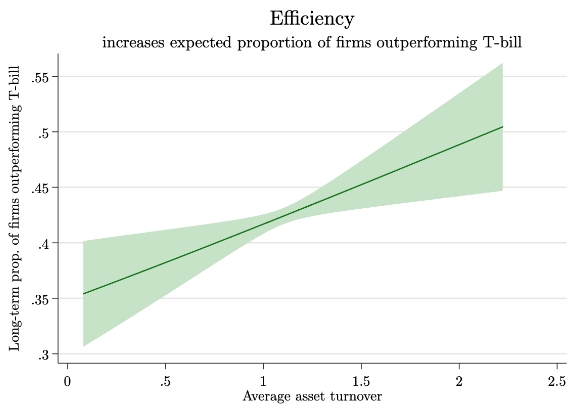
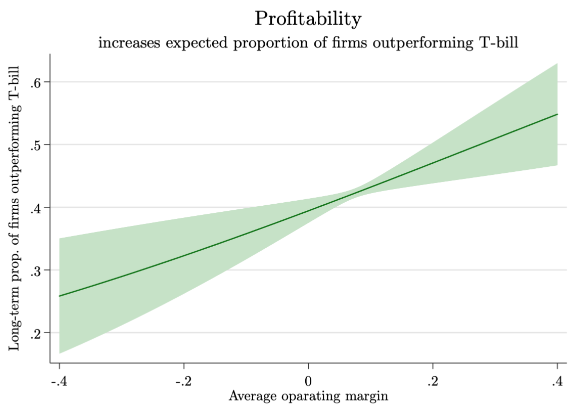



Figure 7: MEP operating margin




Hypothesis 5 states that variability in corporate performance can explain differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset. Regression (4) from Table 15 tests hypothesis 5 and finds some support for it. In the hypothesis development we explain that variability in corporate performance could indicate that firms are more prone to failure. Consequently, we expect variability in corporate performance in a given industry to be associated with a lower proportion of firms that is able to outperform the risk-free rate over their lifetime. However, results from regression (4) and (5) in Table 15 are mixed.

A higher average variability in operating margins within firms of a given industry is associated with more firms outperforming US Treasury bills over their lifetime (t-stat=1.97; t-stat=2.97). This is a counter-intuitive result that could be driven by a number of industries that have very low variability in operating margins and a low proportion of firms that outperform the risk-free rate such as the restaurant and fabricated products industry. At the same time, we observe that industries with very high variability in operating margins such as the precious metals and pharmaceutical industry also have a low proportion of stocks that outperform the risk-free rate. Hence, the relationship between variability in operating margins and outperformance of the risk-free rate could be concave.

A higher variability in average cash flow/debt within firms of a given industry is associated with a lower proportion of stocks outperforming US T-bills (t-stat=-3.43). This is a result that we expected, since a more volatile ability to fulfill long-term obligations could indicate firms in a given industry are more risky and prone to destroy wealth. Furthermore, a higher average variability in interest coverage within firms of a given industry is associated with more stocks outperforming US T-bills in the long run. We disregard this result, because inspection of the marginal effects plot learns that the confidence interval is very wide and the relationship disappears when including other variables from other hypotheses (t-stat=0.46) in regression (5). Average variability in capital ratio (t-stat=0.39), asset turnover (t-stat=1.40), quick ratio (t-stat=-1.25) and return on capital employed (t-stat=-1.48) within firms of a given industry have no significant explanatory power in explaining differences in outperformance.

Regression (5) tests hypothesis 5 in combination with variables from other hypotheses and also finds evidence in support of hypothesis 5. Inclusion of other variables does not alter the result that average variability in operating margin (t-stat=2.97) and cash flow/debt (t-stat=-3.96) within firms of a given industry have explanatory power. Average variability in interest coverage (t-stat=0.46) lost significance in regression (5). However, higher average variability in return on capital employed (t-stat=4.60) within firms of a given industry is associated with more firms outperforming US T-bills over their lifetime. A possible economic explanation is provided by Mauboussin and Callahan (2013) whom show that firms with rapidly increasing returns on capital, i.e. high variability, generate significant returns. To inspect the practical meaning of the results, Figure 8 displays the marginal effects plot of variability in return on capital employed from regression (5). Additionally, Figure 9 displays the relationship between variability in cash flow/debt and the proportion of firms outperforming US T-bills in the long run. Thus, this paper finds evidence in support of hypothesis 5 which states that variability in corporate performance can explain differences in outperformance the risk-free asset across US industries.

Figure 8: MEP ROCE SD

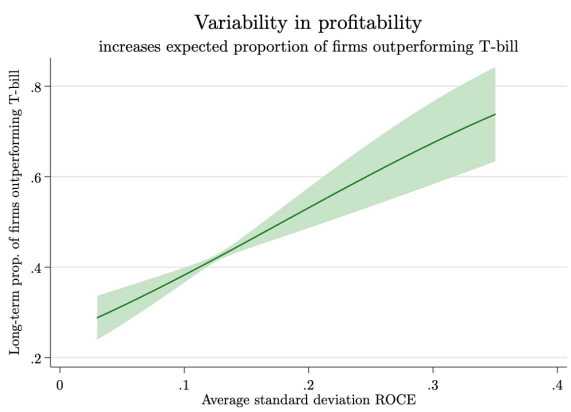
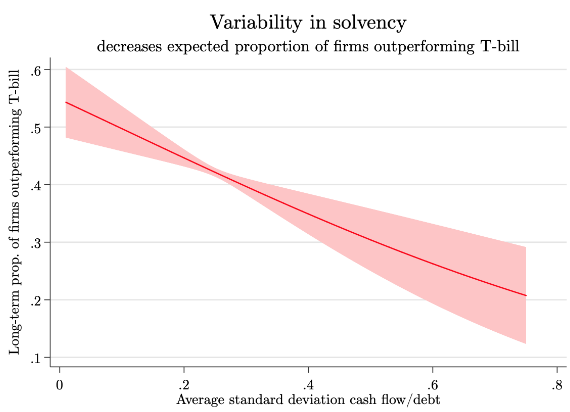



Figure 9: MEP cash/debt SD




In conclusion, substantial evidence is found for hypotheses 3, 4 and 5. Differences in the proportion of stocks across US industries that achieve a higher buy-and-hold return over their lifetime tan the risk-free asset is explained by short-term stock return characteristics, corporate performance of industries and variability in corporate performance within firms. Differences in the proportion of stocks that outperform the risk-free asset are related to differences in the proportion of firms that create or destroy wealth. Hence, determinants for outperforming the risk-free asset might also drive cross-industry variation in concentration of wealth creation. This is tested through hypotheses 7, 8 and 9 in subsequent sections.

### 6.3 Long-term wealth creation by stocks in US industries

The degree to which firms in an industry outperform a given benchmark can obscure the actual experience of investors in aggregate. Wealth creation by stocks can be more informative. From preliminary analysis of Table 16 it becomes clear that some industries create vast quantities of gross wealth whereas others generate almost no wealth. With respect to gross wealth, the computer software, pharmaceutical and petroleum and natural gas industry are the clear winners. They collectively create roughly 22% of total gross wealth from 1926 to 2018. Industries that have not generated significant amounts of gross wealth are the coal, agriculture and textile industry.

When taking into account firms that destroy wealth, absolute amounts of wealth creation decrease substantially. Total net wealth created in the period 1926 to 2018 is $39 trillion whereas gross wealth creation is $ 48 trillion, meaning that value destroying firms destroy roughly $9 trillion in wealth over this period. Additionally, some industries suffer more from value destroying firms than others. For example, gross wealth created from 1926 to 2018 by the electronic equipment industry is roughly $1400 billion whereas net wealth created is $315 billion. This indicates that large wealth destroyers are present in the electronic equipment industry. Notable observations also occur in the fabricated products and coal industry. They have created almost no net wealth and the precious metal industry destroyed wealth over the entire period of 93 years.

It is also interesting to observe that some industries have really taken off in the later part of the sample. For example, the pharmaceutical industry created $2867 billion in net wealth from 1926 until 2018 of which $2641 billion was created from 1970 until 2018. Hence, 92% of net wealth in the pharmaceutical industry was created in the second half of the sample. In various other industries the reverse is visible. For example, from 1926 until 2018 the steel industry created $432 billion in net wealth and from 1970 until 2018 it created $102 billion. Absolute numbers of wealth creation should be interpreted with caution since some industries might be more represented in the stock market or make up a larger portion of the economy. The number of firms listed on the stock market for a particular industry could play a role, especially if large private players exist in the economy. For example, some industries such as the defense industry seem very promising from a buy-and-hold returns perspective with a mean annual return per stock of 22%. However, the net wealth created by this industry is a mere 0.3% of the total wealth created by the market. The main cause is that the number of stocks in the defense industry is limited and their collective market capitalization is low.

By combining absolute figures of wealth creation with stock return benchmarking against Treasury bills and CPI additional insights appear. Combining results from both helps understanding wealth concentration intuitively. For instance, from 1926 to 2018 just 32% of stocks in the computer software industry outperformed the Treasury bill over their lifetime. However, the computer software industry created a vast quantity of wealth, almost $3 trillion. It is clear that a minority of firms exhibits such exceptional growth that they collectively mitigate and dwarf the underperformance of the majority of firms. This begs the question whether within that minority of firms a few firms create all the net wealth. Hypothesis 6 aims to answer that question and is answered in the next section.

Table 16: Gross wealth creation per industry

| | Gross wealth creation (billions of $) | | | |
|---|---|---|---|---|
| | **1926-2018** | | **1970-2018** | |
| | T-bill | CPI | T-bill | CPI |
| **Fama-French 49 Industries** | | | | |
| Agriculture | $45 | $43 | $35 | $36 |
| Food Products | $1067 | $1037 | $850 | $881 |
| Candy & Soda | $733 | $750 | $691 | $719 |
| Beer & Liquor | $204 | $214 | $166 | $189 |
| Tobacco Products | $761 | $745 | $676 | $680 |
| Recreation | $127 | $135 | $118 | $128 |
| Entertainment | $502 | $509 | $473 | $489 |
| Printing & Publishing | $365 | $382 | $351 | $371 |
| Consumer Goods | $1497 | $1591 | $1348 | $1460 |
| Apparel | $250 | $247 | $230 | $232 |
| Healthcare | $340 | $355 | $339 | $354 |
| Medical Equipment | $1092 | $1114 | $1044 | $1079 |
| Pharmaceutical Products | $3178 | $3243 | $2967 | $3092 |
| Chemicals | $1336 | $1262 | $953 | $993 |
| Rubber and Plastic Products | $89 | $89 | $84 | $86 |
| Textiles | $73 | $63 | $44 | $43 |
| Construction Materials | $518 | $498 | $366 | $392 |
| Construction | $123 | $127 | $118 | $123 |
| Steel Works | $540 | $470 | $235 | $260 |
| Fabricated Products | $21 | $20 | $19 | $18 |
| Machinery | $897 | $881 | $741 | $772 |
| Electrical Equipment | $510 | $515 | $406 | $446 |
| Automobiles and Trucks | $1022 | $907 | $465 | $510 |
| Aircraft | $659 | $658 | $608 | $624 |
| Shipbuilding, Railroad Equipment | $91 | $88 | $69 | $72 |
| Defense | $131 | $135 | $131 | $135 |
| Precious Metals | $91 | $92 | $88 | $90 |
| Mines | $249 | $240 | $184 | $185 |
| Coal | $38 | $32 | $17 | $17 |
| Petroleum and Natural Gas | $3772 | $3536 | $2790 | $2846 |
| Utilities | $2007 | $2005 | $1583 | $1698 |
| Communication | $2270 | $2302 | $1973 | $2072 |
| Personal Services | $96 | $101 | $93 | $99 |
| Business Services | $1799 | $1815 | $1774 | $1797 |
| Computers | $2498 | $2565 | $2062 | $2254 |
| Computer Software | $3831 | $3854 | $3831 | $3854 |
| Electronic Equipment | $1384 | $1414 | $1342 | $1373 |
| Measuring and Control Equipment | $589 | $581 | $513 | $529 |
| Business Supplies | $336 | $353 | $289 | $315 |
| Shipping Containers | $264 | $236 | $105 | $127 |
| Transportation | $1017 | $986 | $852 | $866 |
| Wholesale | $842 | $868 | $827 | $858 |
| Retail | $2504 | $2503 | $2216 | $2290 |
| Restaurant, Hotels, Motels | $769 | $783 | $748 | $767 |
| Banking | $2199 | $2240 | $2150 | $2205 |
| Insurance | $2217 | $2274 | $2203 | $2263 |
| Real Estate | $774 | $766 | $769 | $762 |
| Trading | $2503 | $2710 | $2455 | $2675 |
| Other | $99 | $101 | $98 | $100 |
| Not available | $56 | $56 | $47 | $50 |
| Total | $48370 | $48489 | $42536 | $44273 |

The table shows the gross wealth created over two horizons and with respect the risk-free asset (T-bills) and inflation (CPI). Gross wealth creation is calculated by summing all the wealth creation by companies that created positive wealth over their lifetime. All values are in billions of dollar.

Table 17: Net wealth creation per industry

| | Net wealth creation (billions of $) | | | |
|---|---|---|---|---|
| | **1926-2018** | | **1970-2018** | |
| | T-bill | CPI | T-bill | CPI |
| **Fama-French 49 Industries** | | | | |
| Agriculture | $31 | $30 | $21 | $23 |
| Food Products | $999 | $969 | $782 | $812 |
| Candy & Soda | $727 | $745 | $686 | $713 |
| Beer & Liquor | $194 | $206 | $149 | $176 |
| Tobacco Products | $758 | $743 | $674 | $679 |
| Recreation | $64 | $77 | $46 | $63 |
| Entertainment | $423 | $429 | $397 | $412 |
| Printing & Publishing | $308 | $327 | $292 | $315 |
| Consumer Goods | $1392 | $1500 | $1048 | $1247 |
| Apparel | $198 | $200 | $174 | $183 |
| Healthcare | $231 | $249 | $230 | $248 |
| Medical Equipment | $1013 | $1040 | $960 | $1002 |
| Pharmaceutical Products | $2867 | $2917 | $2641 | $2755 |
| Chemicals | $1226 | $1154 | $844 | $887 |
| Rubber and Plastic Products | $53 | $54 | $50 | $52 |
| Textiles | $50 | $43 | $20 | $23 |
| Construction Materials | $455 | $448 | $293 | $336 |
| Construction | $64 | $72 | $65 | $73 |
| Steel Works | $432 | $362 | $102 | $133 |
| Fabricated Products | $2 | $4 | $-0 | $2 |
| Machinery | $546 | $493 | $376 | $372 |
| Electrical Equipment | $281 | $260 | $101 | $138 |
| Automobiles and Trucks | $874 | $762 | $310 | $362 |
| Aircraft | $641 | $644 | $599 | $617 |
| Shipbuilding, Railroad Equipment | $83 | $80 | $62 | $64 |
| Defense | $122 | $127 | $124 | $128 |
| Precious Metals | $-21 | $-21 | $-28 | $-27 |
| Mines | $119 | $107 | $37 | $46 |
| Coal | $7 | $1 | $-10 | $-12 |
| Petroleum and Natural Gas | $3194 | $2960 | $2206 | $2264 |
| Utilities | $1859 | $1846 | $1441 | $1543 |
| Communication | $1460 | $1428 | $1161 | $1195 |
| Personal Services | $32 | $38 | $28 | $36 |
| Business Services | $1243 | $1226 | $1213 | $1204 |
| Computers | $2198 | $2285 | $1710 | $1938 |
| Computer Software | $2956 | $2915 | $2955 | $2914 |
| Electronic Equipment | $315 | $228 | $262 | $186 |
| Measuring and Control Equipment | $531 | $528 | $456 | $477 |
| Business Supplies | $298 | $318 | $250 | $280 |
| Shipping Containers | $238 | $216 | $76 | $104 |
| Transportation | $834 | $823 | $689 | $715 |
| Wholesale | $655 | $686 | $637 | $674 |
| Retail | $2270 | $2283 | $1947 | $2051 |
| Restaurant, Hotels, Motels | $670 | $698 | $648 | $682 |
| Banking | $1726 | $1708 | $1670 | $1667 |
| Insurance | $1822 | $1837 | $1810 | $1827 |
| Real Estate | $632 | $626 | $627 | $621 |
| Trading | $2227 | $2456 | $2189 | $2430 |
| Other | $62 | $66 | $61 | $65 |
| Not available | $32 | $34 | $24 | $29 |
| Total | $39392 | $39224 | $33103 | $34723 |

The table shows the net wealth created over two horizons and with respect to the risk-free asset (T-bill) and inflation (CPI). Net wealth creation is calculated by summing all the wealth created by all stock in the sample over their lifetime. All values are in billions of dollar.

### 6.3.1 Concentration of gross wealth creation in US industries

Creation of gross wealth is concentrated across most industries. Table 18 shows what percentage of gross wealth is created by the top 0.5%, top 1% and top 2.5% of firms in every industry. In some industries the top 1% of firms created more than half of all the gross wealth in that industry. For example, 1,677 firms in the banking industry created positive wealth during their lifetime. However, 17 firms (top 1%) were responsible for 57% of gross wealth creation in the banking industry. Other industries exhibit less concentration. Within the personal services industry the top 1% of firms is responsible for 17% of the gross wealth creation. Regarding gross wealth creation by the top 0.5% of firms, the consumer goods industry was the most concentrated with 68% of gross wealth created by just 2 firms, namely General Electric which had a large division making household appliances and Proctor & Gamble the almost two centuries old consumer goods company. Least concentrated with respect to the 0.5% statistics was the utilities industry, where the top 0.5% of firms created just 9% of gross wealth. This result could be due the fixed infrastructure and regional concentration of utility companies. Both their means of production and distribution system is very capital intensive and often only operated by one company. In the computer industry, the top 1% of firms (Apple, IBM and Cisco Systems) created 83% of gross wealth, which makes it the most concentrated in this category. The agriculture industry is the least concentrated, because just 16% of gross wealth is created by the top 1% of firms. In the top 2.5% category, the computer industry is most concentrated and the coal industry least concentrated.

The top 0.5% statistic is not different from the top 1% statistic in some industries, due to the fact that there are relatively few firms in that industry that created wealth during their lifetime. For example, if in the candy & soda industry 45 firms created wealth, 0.5% would be 0.225 firms and 1% would be 0.45 firms. However, a firm is one economic entity and cannot be divided up in several parts just for statistical convenience. Therefore, the top 0.5% and top 1% are the same for industries with relatively few number of firms. Consequently, in regression analysis we use the proportion of gross wealth created by the top 2.5% of firms per industry as a dependent variable and not the proportion of gross wealth created by the top 0.5% and 1%. Gross wealth concentration levels observed across industries suggest we might fail to reject hypothesis 6, which states that a high degree of concentration in wealth creation, i.e. $< 5\%$ of stocks are responsible for 100% of wealth creation across industries. The next section describes net wealth creation and thereby tests this hypothesis formally.

### 6.3.2 Concentration of net wealth creation in US industries

A similar pattern of wealth concentration occurs in net wealth creation. Table 19 gives an overview of the degree of concentration in net wealth creation across US industries. In a large majority of industries a small percentage of firms is responsible for 100% of the net wealth creation. For example, in the recreation industry, just 1% (Hasbro, Motorala and EMC Corp.) of firms is responsible for 100% of net wealth creation in the period from 1926 to 2018. The tobacco products industry suffered least from concentration in wealth creation: the top 37.5% of firms were responsible for all net wealth creation.

Concentration across all firms regardless of industry is also severe, 4.9% is responsible for all net wealth creation in the period 1926 to 2018. The degree of concentration in net wealth creation is lower than what Bessembinder (2018) finds in a similar period, because this paper allows for industry switching. Hence, some companies are counted double in our statistics. Hypothesis 6 is tested in a single sample t-test to determine whether the proportion of stocks responsible for 100% of net wealth creation in US industries was lower than 0.05. Mean proportion of stocks responsible for 100% of net wealth creation ($0.072 \pm 0.077$) was not significantly lower than 0.05, t(49)=2.0106, p=0.9751[12]. Thus, we reject hypothesis 6.

It is also interesting to compare the total number of firms from the gross wealth creation Table 18 with the net wealth creation Table 19. Large differences exist across industries in the percentage of firms that create wealth. For instance, in the petroleum and natural gas industry just 468 out of 1,471 firms manage to create wealth during their lifetime, which is 32% of the total number of firms in that industry. Comparing the periods 1926-2018 and 1970-2018 in terms of wealth concentration shows that some industries had a higher degree of concentration, some had lower degree of concentration and others had retained a similar degree of concentration. Most industries show a higher degree of concentration in the period from 1970-2018 compared to 1926-2018. For example, in the construction materials industry 12.8% of firms were responsible for all net wealth creation from 1926 to 2018. However, from 1970 to 2018 a significantly lower 7.2% of firms were responsible for all net wealth creation. This could be due to the fact that there was some industry consolidation.

[12] reported p-value evaluates whether proportion was lower than 0.05

Table 18: Concentration of gross wealth creation during 1926-2018

| | Total firms | Top 0.5% of firms | | Top 1% of firms | | Top 2.5% of firms | |
|---|---|---|---|---|---|---|---|
| | # firms | # firms | % of gross wealth | # firms | % of gross wealth | # firms | % of gross wealth |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 43 | 1 | 16 | 1 | 16 | 2 | 33 |
| Food Products | 261 | 2 | 17 | 3 | 22 | 7 | 39 |
| Candy & Soda | 45 | 1 | 47 | 1 | 47 | 2 | 80 |
| Beer & Liquor | 46 | 1 | 34 | 1 | 34 | 2 | 52 |
| Tobacco Products | 24 | 1 | 50 | 1 | 50 | 1 | 50 |
| Recreation | 80 | 1 | 33 | 1 | 33 | 2 | 67 |
| Entertainment | 154 | 1 | 28 | 2 | 51 | 4 | 62 |
| Printing & Publishing | 115 | 1 | 22 | 2 | 41 | 3 | 53 |
| Consumer Goods | 212 | 2 | 68 | 3 | 74 | 6 | 82 |
| Apparel | 128 | 1 | 41 | 2 | 57 | 4 | 68 |
| Healthcare | 244 | 2 | 27 | 3 | 33 | 7 | 47 |
| Medical Equipment | 296 | 2 | 26 | 3 | 34 | 8 | 58 |
| Pharmaceutical Products | 432 | 3 | 29 | 5 | 40 | 11 | 64 |
| Chemicals | 263 | 2 | 35 | 3 | 40 | 7 | 56 |
| Rubber and Plastic Products | 113 | 1 | 23 | 2 | 33 | 3 | 42 |
| Textiles | 104 | 1 | 17 | 2 | 31 | 3 | 39 |
| Construction Materials | 360 | 2 | 20 | 4 | 28 | 10 | 42 |
| Construction | 127 | 1 | 12 | 2 | 22 | 4 | 40 |
| Steel Works | 206 | 2 | 16 | 3 | 24 | 6 | 42 |
| Fabricated Products | 44 | 1 | 17 | 1 | 17 | 2 | 32 |
| Machinery | 414 | 3 | 28 | 5 | 34 | 11 | 49 |
| Electrical Equipment | 245 | 2 | 29 | 3 | 37 | 7 | 54 |
| Automobiles and Trucks | 168 | 1 | 43 | 2 | 53 | 5 | 66 |
| Aircraft | 76 | 1 | 38 | 1 | 38 | 2 | 57 |
| Shipbuilding, Railroad Equipment | 36 | 1 | 50 | 1 | 50 | 1 | 50 |
| Defense | 25 | 1 | 58 | 1 | 58 | 1 | 58 |
| Precious Metals | 79 | 1 | 16 | 1 | 16 | 2 | 26 |
| Mines | 110 | 1 | 14 | 2 | 25 | 3 | 34 |
| Coal | 37 | 1 | 21 | 1 | 21 | 1 | 21 |
| Petroleum and Natural Gas | 468 | 3 | 41 | 5 | 50 | 12 | 66 |
| Utilities | 302 | 2 | 9 | 4 | 15 | 8 | 26 |
| Communication | 329 | 2 | 20 | 4 | 30 | 9 | 48 |
| Personal Services | 134 | 1 | 11 | 2 | 17 | 4 | 27 |
| Business Services | 757 | 4 | 33 | 8 | 48 | 19 | 59 |
| Computers | 205 | 2 | 62 | 3 | 83 | 6 | 88 |
| Computer Software | 562 | 3 | 57 | 6 | 72 | 15 | 81 |
| Electronic Equipment | 429 | 3 | 38 | 5 | 47 | 11 | 61 |
| Measuring and Control Equipment | 206 | 2 | 28 | 3 | 37 | 6 | 56 |
| Business Supplies | 148 | 1 | 23 | 2 | 29 | 4 | 40 |
| Shipping Containers | 103 | 1 | 39 | 2 | 45 | 3 | 51 |
| Transportation | 315 | 2 | 20 | 4 | 30 | 8 | 46 |
| Wholesale | 482 | 3 | 19 | 5 | 26 | 13 | 44 |
| Retail | 550 | 3 | 35 | 6 | 45 | 14 | 60 |
| Restaurant, Hotels, Motels | 212 | 2 | 43 | 3 | 50 | 6 | 61 |
| Banking | 1,677 | 9 | 47 | 17 | 57 | 42 | 72 |
| Insurance | 407 | 3 | 35 | 5 | 41 | 11 | 54 |
| Real Estate | 450 | 3 | 19 | 5 | 24 | 12 | 38 |
| Trading | 1,619 | 9 | 29 | 17 | 40 | 41 | 58 |
| Other | 53 | 1 | 26 | 1 | 26 | 2 | 47 |
| Not available | 220 | 2 | 33 | 3 | 41 | 6 | 59 |
| Total | 14115 | 102 | 28 | 172 | 37 | 379 | 53 |

The table shows how concentrated gross wealth creation is per industry. Firms are allowed to switch industries. Percentages show how much of the total gross wealth was created by the top x% of stocks.

Table 19: Concentration of net wealth creation, based on T-bill

| | 1926-2018 | | | 1970-2018 | | |
|---|---|---|---|---|---|---|
| | # total firms | Created 100% net wealth | | # total firms | Created 100% net wealth | |
| **Fama-French 49 Industries** | | | | | | |
| Agriculture | 133 | 6 | 4.5% | 128 | 5 | 3.9% |
| Food Products | 461 | 71 | 15.4% | 377 | 44 | 11.7% |
| Candy & Soda | 73 | 17 | 23.3% | 67 | 16 | 23.9% |
| Beer & Liquor | 95 | 12 | 12.6% | 84 | 7 | 8.3% |
| Tobacco Products | 40 | 15 | 37.5% | 30 | 13 | 43.3% |
| Recreation | 309 | 3 | 1.0% | 291 | 2 | 0.7% |
| Entertainment | 513 | 21 | 4.1% | 493 | 16 | 3.2% |
| Printing & Publishing | 249 | 17 | 6.8% | 240 | 14 | 5.8% |
| Consumer Goods | 535 | 20 | 3.7% | 489 | 4 | 0.8% |
| Apparel | 335 | 11 | 3.3% | 302 | 6 | 2.0% |
| Healthcare | 646 | 18 | 2.8% | 646 | 18 | 2.8% |
| Medical Equipment | 797 | 55 | 6.9% | 793 | 50 | 6.3% |
| Pharmaceutical Products | 1,293 | 57 | 4.4% | 1,282 | 50 | 3.9% |
| Chemicals | 474 | 61 | 12.9% | 428 | 44 | 10.3% |
| Rubber and Plastic Products | 251 | 7 | 2.8% | 240 | 7 | 2.9% |
| Textiles | 241 | 14 | 5.8% | 200 | 3 | 1.5% |
| Construction Materials | 690 | 88 | 12.8% | 607 | 44 | 7.2% |
| Construction | 379 | 7 | 1.8% | 361 | 7 | 1.9% |
| Steel Works | 396 | 30 | 7.6% | 315 | 4 | 1.3% |
| Fabricated Products | 115 | 1 | 0.9% | 115 | 1 | 0.9% |
| Machinery | 843 | 22 | 2.6% | 751 | 12 | 1.6% |
| Electrical Equipment | 812 | 8 | 1.0% | 736 | 3 | 0.4% |
| Automobiles and Trucks | 388 | 17 | 4.4% | 318 | 9 | 2.8% |
| Aircraft | 134 | 24 | 17.9% | 111 | 21 | 18.9% |
| Shipbuilding, Railroad Equipment | 64 | 10 | 15.6% | 50 | 6 | 12.0% |
| Defense | 47 | 7 | 14.9% | 47 | 8 | 17.0% |
| Precious Metals | 341 | 3 | 0.9% | 338 | 4 | 1.2% |
| Mines | 330 | 6 | 1.8% | 290 | 3 | 1.0% |
| Coal | 76 | 1 | 1.3% | 56 | 3 | 5.4% |
| Petroleum and Natural Gas | 1,471 | 38 | 2.6% | 1,384 | 26 | 1.9% |
| Utilities | 414 | 123 | 29.7% | 381 | 107 | 28.1% |
| Communication | 831 | 18 | 2.2% | 817 | 17 | 2.1% |
| Personal Services | 388 | 7 | 1.8% | 381 | 6 | 1.6% |
| Business Services | 2,095 | 37 | 1.8% | 2,066 | 34 | 1.6% |
| Computers | 826 | 7 | 0.8% | 815 | 5 | 0.6% |
| Computer Software | 1,878 | 11 | 0.6% | 1,878 | 10 | 0.5% |
| Electronic Equipment | 1,281 | 1 | 0.1% | 1,246 | 1 | 0.1% |
| Measuring and Control Equipment | 544 | 34 | 6.2% | 522 | 28 | 5.4% |
| Business Supplies | 239 | 39 | 16.3% | 215 | 31 | 14.4% |
| Shipping Containers | 169 | 25 | 14.8% | 137 | 7 | 5.1% |
| Transportation | 723 | 40 | 5.5% | 617 | 29 | 4.7% |
| Wholesale | 1,329 | 55 | 4.1% | 1,288 | 51 | 4.0% |
| Retail | 1,373 | 90 | 6.6% | 1,296 | 64 | 4.9% |
| Restaurant, Hotels, Motels | 599 | 32 | 5.3% | 585 | 28 | 4.8% |
| Banking | 2,538 | 63 | 2.5% | 2,499 | 56 | 2.2% |
| Insurance | 714 | 50 | 7.0% | 705 | 48 | 6.8% |
| Real Estate | 919 | 84 | 9.1% | 896 | 81 | 9.0% |
| Trading | 2,638 | 210 | 8.0% | 2,587 | 206 | 8.0% |
| Other | 149 | 4 | 2.7% | 147 | 4 | 2.7% |
| Not available | 525 | 6 | 1.1% | 513 | 4 | 0.8% |
| Total | 32,703 | 1603 | 4.9% | 31,160 | 1266 | 4.1% |

The table shows how concentrated net wealth creation is per industry. Percentages in the columns show what percentage of firms was responsible for 100% of the net wealth creation.

## 6.4 Assessing cross-industry variation in concentration of gross wealth creation

Table 20 reports results of testing hypotheses 7, 8 and 9. Hypothesis 7 claims that short-term return volatility, return skewness and mean returns can explain differences in the proportion of stocks that is responsible for a x% of wealth across US industries. Regressions (1) and (2) from Table 20 regress average monthly standard deviation of returns, return skewness and mean returns of a given industry on the proportion of gross wealth created by the top 2.5% of firms. Regression (5) does the same, but includes variables from hypotheses 8 and 9 as controls.

In the period 1926 to 2018, average short-term volatility is positively associated with the proportion of wealth created by the top 2.5% performing stocks in each industry (t-stat=2.23). This result is in line with the theory proposed by Farago and Hjalmarsson (2019) and Bessembinder et al. (2019) that short-term return volatility drives extreme positive skewness in long-term stock returns. From an economic standpoint we could argue that higher short-term volatility is associated with a riskier business environment where only a few firms survive thereby inducing concentration in wealth creation. However, in the period 1970 to 2018 average short-term volatility bears no significant explanatory power (t-stat=1.74; t-stat=0.69). Average skewness in monthly returns is negatively associated (t-stat=-2.19) with concentration of wealth creation in the period 1926 to 2018. However, subsequent regressions find that skewness does not have significant explanatory power for concentration of wealth creation (t-stat=-1.29; t-stat=-1.15). Furthermore, mean stock returns do not bear significant explanatory power in explaining the proportion of gross wealth created by the top 2.5% performing firms in each industry (t-stat=1.41; t-stat=1.05). However, in regression (5) a higher mean stock return is associated with a higher degree of concentration in gross wealth creation (t-stat=3.61). Figure 10 depicts that relationship. Therefore, evidence for hypothesis 7 is mixed and we are inclined to reject it.

Figure 10: MEP mean return


Figure 11: MEP capital ratio

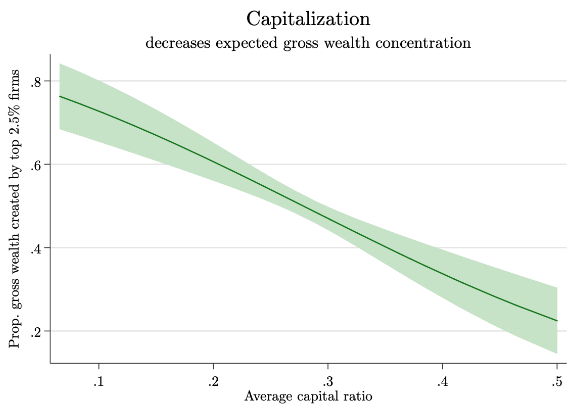


Table 20: Cross-industry gross wealth concentration by top 2.5% of firms, based on T-bill

| | Proportion of gross wealth created by top 2.5% of firms | | | | |
|---|---|---|---|---|---|
| | 1926-2018 | 1970-2018 | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Standard deviation | 8.149* | 9.067 | | | 3.058 |
| | (2.23) | (1.74) | | | (0.69) |
| Skewness | −1.437* | −1.323 | | | −0.752 |
| | (−2.19) | (−1.29) | | | (−1.15) |
| Mean return | 19.41 | 18.41 | | | 54.37*** |
| | (1.41) | (1.05) | | | (3.61) |
| Capital ratio | | | −4.388*** | | −5.985*** |
| | | | (−3.38) | | (−4.97) |
| Asset turnover | | | −0.114 | | −0.187 |
| | | | (−0.36) | | (−0.49) |
| Quick ratio | | | −0.147 | | −0.841* |
| | | | (−0.40) | | (−2.37) |
| ROCE | | | 13.73*** | | 12.64*** |
| | | | (3.88) | | (3.63) |
| Operating margin | | | 0.216 | | −1.882 |
| | | | (0.12) | | (−0.86) |
| Cash flow/debt | | | −2.800 | | −0.937 |
| | | | (−1.42) | | (−0.58) |
| Interest coverage | | | −0.209** | | −0.286*** |
| | | | (−2.72) | | (−4.86) |
| Capital ratio SD | | | | 3.038 | 6.149 |
| | | | | (0.46) | (1.19) |
| Asset turnover SD | | | | −3.525 | −6.498* |
| | | | | (−1.62) | (−2.26) |
| Quick ratio SD | | | | 0.236 | 0.418 |
| | | | | (0.44) | (0.74) |
| ROCE SD | | | | 7.369 | 14.06** |
| | | | | (1.91) | (2.80) |
| Operating margin SD | | | | −1.720** | −2.274*** |
| | | | | (−2.94) | (−3.43) |
| Cash flow/debt SD | | | | 1.695 | 2.230 |
| | | | | (0.84) | (1.23) |
| Interest coverage SD | | | | 0.00748 | −0.0260 |
| | | | | (0.54) | (−1.59) |
| Constant | −0.114 | −0.509 | 1.162 | −0.655 | 2.563** |
| | (−0.19) | (−0.85) | (1.09) | (−0.97) | (2.78) |
| Observations | 50 | 50 | 49 | 50 | 49 |

*t* statistics in parentheses
This fractional logistic regression measures the effect of the short-term mean, standard deviation (SD) and skewsness of stock returns on long-term concentration of gross wealth creation across industries. The mean, SD and skewness are calculated for each stock and averaged to get to the industry figures. In regressions (3), (4) and (5) corporate performance figures and the standard deviation of those figures is also taken into account.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Hypothesis 8 states that corporate performance can explain differences in the proportion of stocks that is responsible for a x% of wealth across US industries. Regressions (3) and (5) in Table 20 formally test this hypothesis and find evidence for it. We expect better corporate performance in a given industry to be associated with less concentration in wealth creation, because if firms as a group perform better it is likely that wealth creation is more evenly spread.

First, in both regressions (3) and (5) in Table 20 a higher average capital ratio, i.e. more leverage, in a given industry is associated with a lower degree of concentration in wealth creation (t-stat=-3.38; t-stat=-4.97), which is also displayed in Figure 11. This result is consistent with findings from evaluating cross-industry variation in outperformance relative to T-bills which are reported in section 6.2. In that section we find that a higher average capital ratio in a given industry is associated with a larger proportion of firms being able to create wealth during their lifetime. Industries where firms are more leveraged could be more regulated, such as the banking, utilities and tobacco industry (all have relatively high leverage and low concentration of wealth creation), which would induce a more even distribution of wealth creation.

Second, a higher average return on capital employed in a given industry is positively associated (t-stat=3.88; t-stat=3.63) with the proportion of gross wealth created by top performing firms. Hence, more profitable industries often suffer from more concentration in wealth creation. Figure 12 displays the relationship between return on capital employed and concentration of gross wealth creation. This is against our expectations as we would expect industries where firms are on average more profitable to have a more even distribution of wealth creation. Theoretically, firms in very profitable industries might have more resources to mitigate competition by increasing barriers to entry through substantial spending on R&D or marketing (Porter (1980)).

Third, a higher average interest coverage in a given industry is associated (t-stat=-2.72;t-stat=-4.86) with a lower degree of concentration in wealth creation. Figure 13 plots the marginal effects of average interest coverage. This is a resulted we expected in the hypothesis development, because a better ability to service interest payments lowers the probability of default as Altman (1968) found. Hence, if there are less value destroying bankrupt firms in an industry it decreases the concentration in gross wealth creation.

Last, a higher average quick ratio (better balance sheet liquidity) is negatively associated (t-stat=-2.37) with the proportion of wealth ascribed to the top performing firms in each industry. This is also a result that we expected since better liquidity indicates less financial distress among firms, which tends to increase returns as Piotroski (2000) found. Other corporate performance measures such as asset turnover, operating margin and cash/flow debt are not significant predictors for concentration of wealth creation across industries. However, capital ratio, ROCE and interest coverage were highly significant ($p < 0.001$) and remained significant when controlling for other variables. Therefore, we determine that evidence in support of hypothesis 8 is convincing. However, better corporate performance does not necessarily imply that industries are less concentrated.

Figure 12: MEP ROCE

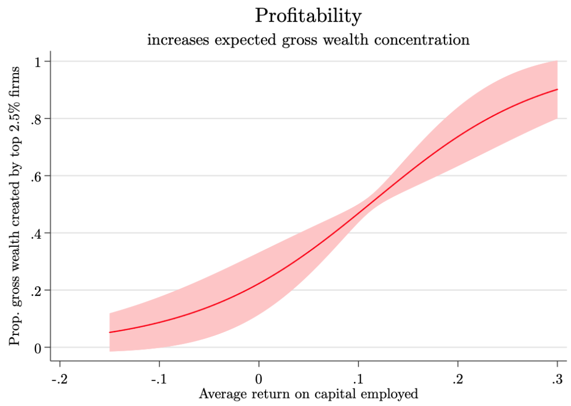
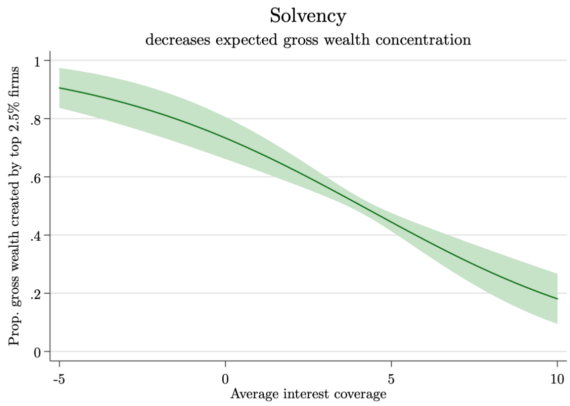



Figure 13: MEP interest coverage




Hypothesis 9 articulates that variability in corporate performance can explain differences in the proportion of stocks that is responsible for a x% of wealth creation. Regressions (4) and (5) in Table 20 regress average standard deviations of corporate performance measures within firms of a given industry on the proportion of gross wealth created by the top 2.5% performing firms. In regression (4) limited evidence is found for hypothesis 9, because all variables are insignificant. Except for average variability in operating margins within firms of a given industry which is negatively associated (t-stat=-2.94) with concentration in gross wealth creation. This result could indicate that industries where profitability is more volatile are more competitive and thus experience less concentration in gross wealth creation.

Figure 14: MEP asset turnover SD

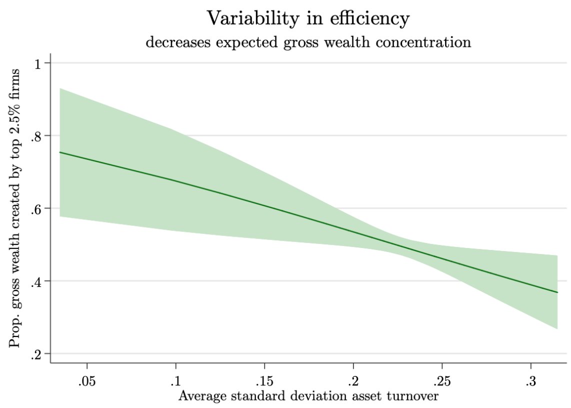
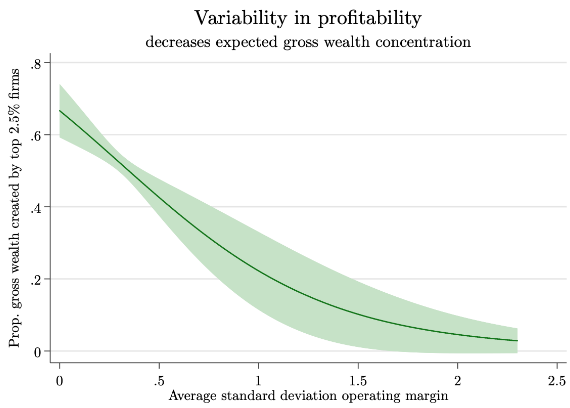



Figure 15: MEP operating margin SD




Stronger evidence in support for hypothesis 9 is found in regression (5) from Table 20. First, a higher average variability in asset turnover is associated (t-stat=-2.26) with a lower proportion of gross wealth produced by the top 2.5% performing firms. Figure 14 displays this relationship. This relationship might be driven by cyclical industries in our sample such as shipbuilding, retail and wholesale. It could be the case that firms in these industries have a relatively low ability to build stable dominant businesses that capture the majority of wealth creation.

Second, a higher average standard deviation of return on capital employed within firms in a given industry is associated (t-stat=2.80) with a higher degree of concentration in gross wealth creation. This result is troublesome, because when assessing cross-industry variation in performance relative to T-bills we find that variability in return on capital employed is associated with more firms being able to create wealth during their lifetime. Hence, the relationship between concentration of wealth creation and an industries' ability to outperform the risk-free rate could be more complicated than initially thought.

Third, average variability in operating margins is negatively associated (t-stat=-3.43) with concentration in gross wealth creation, Figure 15 displays this relationship. Therefore, variability in profitability is not unanimously related to a lower concentration in wealth creation. One potential explanation for this result is that variability in a firm's profitability could increase uncertainty among investors. Pástor and Pietro (2003) point out that uncertainty around profitability increases stock return volatility which potentially induces skewness and concentration in wealth creation. In conclusion, this paper finds mixed evidence in support of hypothesis 9.

### 6.5 Assessing cross-industry variation in concentration of net wealth creation

This section also assesses cross-industry variation in concentration of wealth creation. However, instead of concentration in **gross** wealth creation we use concentration in **net** wealth. Gross wealth excludes value destroying firms whereas net wealth includes all firms. Hypotheses 7, 8 and 9 are tested with fractional outcome regressions where the proportion of firms responsible for all net wealth creation in a given industry is the dependent variable. Hence, significant coefficients with a positive sign are associated with a lower degree of concentration and significant coefficients with a negative sign are associated with a higher degree of concentration. Results are reported in Table 21 and are explained in subsequent paragraphs.

Hypothesis 7 states that short-term stock return volatility, skewness and means can explain differences in the proportion of stocks that is responsible for a x% of wealth across US industries. Regressions (1), (2) and (5) in Table 21 test this hypothesis. First, a higher average monthly return volatility in a given industry is associated (t-stat=-3.73; t-stat=-9.26) with a lower proportion of firms responsible for all net wealth creation. This relationship is displayed in Figure 16. That relationship must be judged with discretion since a positive association is not found (t-stat=1.44) in regression (2). Second, average monthly return skewness is not a significant predictor (t-stat=0.91, t-stat=-1.06; t-stat=-0.56) for concentration in net wealth creation in each industry. Third, higher mean monthly stock returns are associated (t-stat=4.26; t-stat=2.81) with a larger degree of concentration in wealth creation in regressions without control variables. Including control variables in regression (5) renders the coefficient of mean returns insignificant (t-stat=1.85). Hence, short-term return characteristics can partly explain differences in concentration of wealth creation. Thus, this paper interprets hypothesis 7 as partly accepted.

Hypothesis 8 claims that corporate performance can explain differences in the proportion of stocks that is responsible for x% of wealth across US industries. Regressions (3) and (5) in Table 21 test this hypothesis and find inadequate evidence for it. In the stand-alone model (3), both average asset turnover (t-stat=-2.71) and average operating margin (t-stat=2.38) in a given industry are positively associated with concentration in wealth creation. However, model

Figure 16: MEP volatility

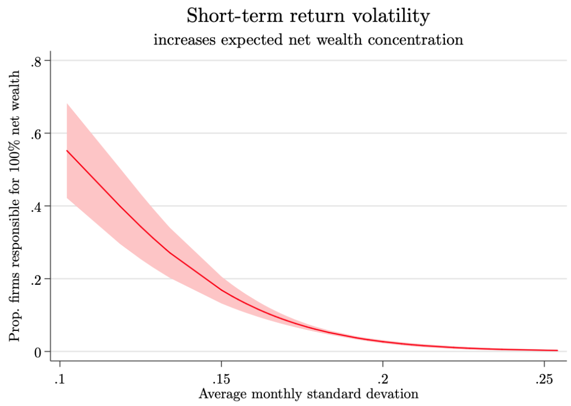
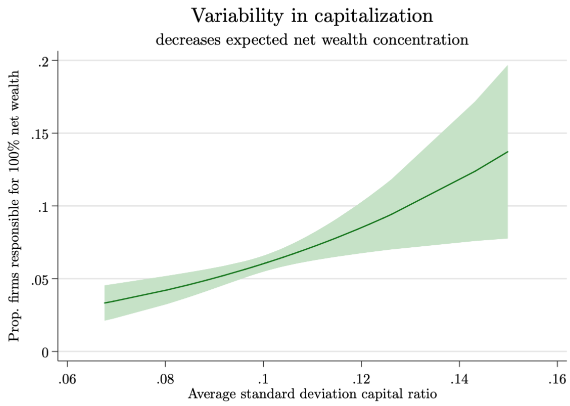



Figure 17: MEP capital ratio SD




(5) includes control variables from the other hypotheses 7 and 9 causing none of the corporate performance variables to have explanatory power. Therefore, corporate performance cannot explain differences in cross-industry variation of concentration in net wealth creation. For that reason, we reject hypothesis 8 with respect to concentration in net wealth creation.

Hypothesis 9 sets forth that variability in corporate performance can explain differences in the proportion of stocks that is responsible for a x% of wealth across US industries. Regression (4) and (5), reported in Table 21 test this hypothesis. Results are moderately consistent with the analysis of cross-industry variation in concentration of gross wealth creation. For example, a higher variability in operating margins within firms of a given industry is associated (t-stat=2.55; t-stat=2.71) with a lower degree of concentration in wealth creation just as in the previous analysis of gross wealth concentration. This is in line with the theory that more competitive industries have a higher variability in profitability which can result in lower industry concentration (Gaspar and Massa (2006)).

Additionally, a higher variability of return on capital employed within firms of a given industry is associated (t-stat=-3.00) with a larger degree of concentration in wealth creation. However, that result becomes insignificant when including control variables. Notably, regression (5) in Table 21, which tests the model including control variables implies that a higher variability in capital ratios (t-stat=3.38) and a higher variability in cash flow/debt (t-stat=3.33) is associated with a larger proportion of firms being responsible for all net wealth creation. This results is in line with the theory by Gaspar and Massa (2006) that more variability in corporate performance induces competition which could lower industry concentration. In conclusion, three out of six variables are highly significant $p < 0.01$ in model (5) with control variables. Thus, we fail to reject hypothesis 9 and find evidence in support of it.

Table 21: Cross-industry net wealth concentration, based on T-bill

| | Proportion of firms responsible for 100% of net wealth creation | | | | |
|---|---|---|---|---|---|
| | 1926-2018 | 1970-2018 | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Standard deviation | $-20.52^{***}$ | $-10.61$ | | | $-43.83^{***}$ |
| | $(-3.73)$ | $(-1.44)$ | | | $(-9.26)$ |
| Skewness | 0.943 | $-1.355$ | | | $-0.325$ |
| | (0.91) | $(-1.06)$ | | | $(-0.56)$ |
| Mean return | $60.88^{***}$ | $70.61^{**}$ | | | 39.29 |
| | (4.26) | (2.81) | | | (1.85) |
| Capital ratio | | | 2.941 | | 1.463 |
| | | | (1.08) | | (0.89) |
| Asset turnover | | | $-1.092^{**}$ | | 0.0418 |
| | | | $(-2.71)$ | | (0.08) |
| Quick ratio | | | $-0.884$ | | $-0.628$ |
| | | | $(-1.19)$ | | $(-1.43)$ |
| ROCE | | | $16.54^{**}$ | | 3.463 |
| | | | (2.84) | | (0.70) |
| Operating margin | | | $-6.193^{*}$ | | 1.305 |
| | | | $(-2.38)$ | | (0.63) |
| Cash flow/debt | | | $-4.000$ | | 0.483 |
| | | | $(-1.50)$ | | (0.25) |
| Interest coverage | | | 0.0439 | | 0.0244 |
| | | | (0.28) | | (0.23) |
| Capital ratio SD | | | | 25.06 | $21.08^{***}$ |
| | | | | (1.46) | (3.38) |
| Asset turnover SD | | | | $-3.223$ | 2.067 |
| | | | | $(-0.76)$ | (0.66) |
| Quick ratio SD | | | | $-4.089^{**}$ | $-1.076$ |
| | | | | $(-2.80)$ | $(-1.42)$ |
| ROCE SD | | | | $-26.17^{**}$ | 4.882 |
| | | | | $(-3.00)$ | (0.87) |
| Operating margin SD | | | | $3.884^{*}$ | $1.595^{**}$ |
| | | | | (2.55) | (2.71) |
| Cash flow/debt SD | | | | $-1.066$ | $5.991^{***}$ |
| | | | | $(-0.24)$ | (3.33) |
| Interest coverage SD | | | | $0.0683^{**}$ | 0.0117 |
| | | | | (3.03) | (0.69) |
| Constant | $-0.210$ | $-0.0197$ | $-2.489$ | $-1.219$ | $-0.134$ |
| | $(-0.27)$ | $(-0.03)$ | $(-1.22)$ | $(-0.67)$ | $(-0.13)$ |
| Observations | 50 | 50 | 49 | 50 | 49 |

*t* statistics in parentheses

This fractional logistic regression measures the effect of the short-term mean, standard deviation (SD) and skewsness of stock returns on long-term concentration of net wealth creation across industries. The mean, SD and skewness are calculated for each stock and averaged to get to industry figures. In regressions (3), (4) and (5) corporate performance figures and the standard deviation of those figures is also taken into account.

$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

In conclusion, mixed evidence is found for hypothesis 5, we reject hypothesis 7 and support hypothesis 9. When combining the results from both the cross-industry analysis of concentration in gross and net wealth concentration this paper fails to reject hypotheses 5, 7 and 9 with discretion. Hence, short-term return characteristics, corporate performance of industries and variability in corporate performance within firms are able to partly explain differences in concentration of wealth creation.

# 7 Robustness checks

It can be discussed what the definition of wealth creation actually is. An investor choosing between a risky asset and an assumed to be risk-free asset defines wealth creation as the excess return achieved by holding the risky asset relative to the risk-free asset. However, a regular person using proceeds from the risky asset to fund real life expenses might be more interested in the excess return above the rate of inflation. Hence, a measure of wealth creation based on inflation might be more suitable. In an economic environment with lower risk-free rates than the the rate of inflation, a phenomena that has occurred in the past decade, wealth creation calculated with the risk-free rate could be artificially high. Hence, calculating wealth creation based on inflation is important. Therefore, this paper calculates all the previously mentioned results while using wealth creation based on the rate of inflation.

The method for calculating wealth creation over the rate of inflation is relatively simple. The equations from section 5.2 need to be altered as if the investor has the opportunity to invest at the rate of inflation. Some real world examples exists of products offering inflation protection, such as ETFs containing inflation-linked government bonds[13]. Another option available to investors is Treasury Inflation Protected Securities (TIPS)[14]. Exactly mimicking a monthly rate of inflation is quite hard as inflation-linked bonds generally have minimum maturities of 5 years and pay coupon semi-annually. However, with some financial engineering it should be possible.

The future value factor equation (13) can be altered by substituting the Treasury bill return $R_f$ by the "return" of inflation $R_i$. In this paper, changes in the consumer price index (hereinafter CPI) are used as a measure for inflation or deflation.

$$FV_{t,T} = (1 + R_{it+1}) * (1 + R_{it+2}) * (1 + R_{it+3}) * ... * (1 + R_{iT}) \tag{34}$$

In the same fashion equation (14) changes by replacing the Treasury bill return $R_f$ with the "return" of inflation $R_i$.

$$\begin{aligned} W_T - W_0 * FV_{0,T} &= I_0 * (R_1 - R_{i1}) * FV_{1,T} \\ &+ I_1 * (R_2 - R_{i2}) * FV_{2,T} + ... \\ &+ I_{T-2} * (R_{T-1} - R_{iT-1}) * FV_{T-1,T} \\ &+ I_{T-1} * (R_T - R_{iT}) \end{aligned} \tag{35}$$

[13] https://www.lyxoretf.nl/pdfDocuments/DTP109998%20-%20Lyxor%20Inflation%20Guide.pdf
[14] https://www.treasurydirect.gov/indiv/products/prod_tips_glance.htm

It is disproportionate to repeat the full empirical results section by reporting and describing results obtained of testing all hypotheses again with the method of calculating wealth creation based on inflation. Therefore subsequent sections briefly summarize whether conclusions from the empirical results section change.

### 7.1 Checking robustness: hypothesis 2

Hypothesis 2 states that the majority of stocks across US industries are unable to achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset. However, wealth creation is not necessarily defined by achieving a return greater than the risk-free rate for all investors. For example, pension funds could define wealth creation for their pensioners by the ability to increase pensions more than the rate of inflation, thereby increasing the purchasing power of pensioners. Therefore, for robustness this paper measures whether individual stocks are able to outperform inflation over their lifetime. Table 13 reports that 44% of stocks in the market are able to achieve a higher buy-and-hold return than the rate of inflation over their lifetime. Percentages differ between industries. Formally testing hypothesis 2 in the context of inflation using a t-test shows that we fail to reject it. Mean proportion of stocks outperforming inflation over their lifetime ($0.46 \pm 0.12$) in a given industry is significantly lower than 0.50, t(49)=-2.4041, p=0.9847.

### 7.2 Checking robustness: hypotheses 3, 4 and 5

When testing hypotheses 3, 4 and 5 for robustness, the dependent variable in the fractional outcome regressions changes. Instead of using the proportion of stocks that is able to outperform US T-bills over their lifetime in a given industry, robustness tests use the proportion of stocks that is able to outperform inflation. Table 22 reports results of the regression testing hypotheses 3, 4 and 5. Significance and magnitude of coefficients is similar to results obtained in the empirical results section. Regarding hypothesis 3 for example, a higher average standard deviation of monthly returns in a given industry is associated (t-stat=-8.49; t-stat=-6.32; t-stat=-9.43) with a lower proportion of firms achieving a higher return than the inflation benchmark. This relationship is displayed in Figure 18. Additionally, with respect to hypothesis 5, a higher average variability in return on capital employed within firms of a given industry is associated (t-stat=4.81) with a higher proportion of firms outperforming inflation over the long run. Figure 19 displays this relationship. This is a result similar to the result in the empirical results section. Therefore, conclusions regarding hypotheses 3, 4 and 5 remain the same. Namely, substantial evidence is found for each of them.

### 7.3 Checking robustness: hypothesis 6

Hypothesis 6 claims that a high degree of concentration in wealth creation is present across US industries. Table 17 shows that there are relatively small differences in wealth creation when measured based on the risk-free asset or inflation. Therefore, we expect to reject hypothesis 3 again. Most industries do suffer from concentration in wealth creation. However, the criterion set by hypothesis 6 is strict, i.e. $< 5\%$ of firms are responsible for all net wealth creation. Hypothesis

Figure 18: MEP volatility

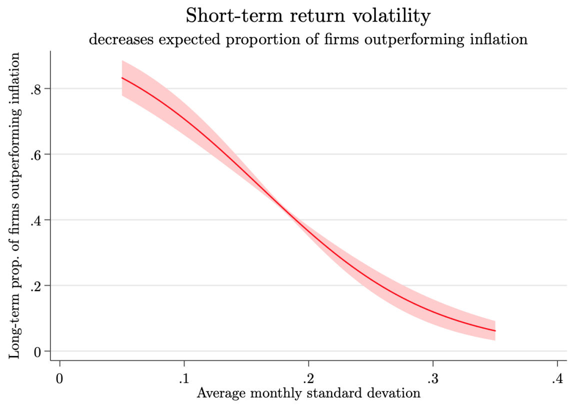
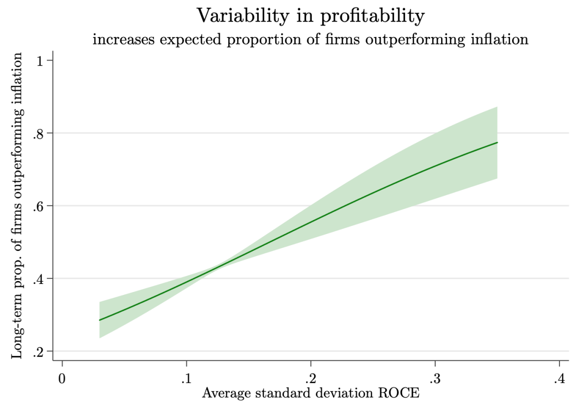



Figure 19: MEP ROCE SD




6 is tested again in a one-sample t-test to determine whether the proportion of stocks responsible for 100% of net wealth creation in US industries was lower than 0.05. Mean proportion of stocks responsible for 100% of net wealth creation ($0.075 \pm 0.078$) is not significantly lower than 0.05, t(49)=2.2254, p=0.9847[15]. Thus, we reject hypothesis 6 again.

### 7.4 Checking robustness: hypotheses 7, 8 and 9

Assessing robustness with respect to hypotheses 7, 8 and 9 is accomplished by re-calculating concentration of wealth creation by using inflation and using outcome of those calculations as the dependent variable in regressions. Table 23 test hypotheses 7, 8 and 9 using concentration in gross wealth creation whereas Table 24 test them using concentration in net wealth creation. These hypotheses claim that short-term return characteristics, corporate performance and variability in corporate performance can explain differences across industries in concentration of wealth creation.

Evidence in support of hypothesis 7 remains mixed, as most short-term return characteristic variable remain insignificant predictors for concentration of in gross wealth creation. However, short-term return characteristics do have significant explanatory power in regression based on concentration in net wealth creation. For example, regression (5) in Table 24 shows that average short-term volatility is associated (t-stat=-9.62) with a higher degree of concentration in wealth creation.

Conclusions for hypothesis 8, that claims corporate performance can explain differences in concentration of wealth creation, remain conflicting. Regressions based on concentration in gross wealth creation find clues in support of hypothesis 8 while regression using concentration of net wealth creation point to rejection of the hypothesis. For example, regarding concentration in gross wealth creation, Figure 20 displays a strong negative association (t-stat=-5.07) between average interest coverage in a given industry and concentration in gross wealth creation. Additionally, Figure 21 displays a strong positive association (t-stat=4.00) between average return on capital employed of a given industry and concentration in gross wealth creation. On the

[15] reported p-value evaluates whether proportion is lower than 0.05

contrary, regression (5) in Table 24 shows that all variables relating corporate performance to concentration in wealth creation have coefficient indistinguishable from zero. Thus, this paper continues to find contradictory evidence for hypothesis 8.

Robustness checks both strengthen and weaken conclusions from the empirical results section that hypothesis 9 can be accepted. Regressions (4) and (5) in Table 23 and Table 24 test hypothesis 9 which states that variability in corporate performance can explain differences in the proportion of stocks that is responsible for a x% of wealth creation across US industries. Results from Table 23 strengthen the conclusion that hypothesis 9 can be accepted. For example, higher average variability in interest coverage within firms of a given industry is negatively associated (t-stat=-2.04) with concentration in gross wealth creation based on inflation. This was not the case when concentration in gross wealth creation was calculated based on US Treasury bills. Hence, more variables contain significant explanatory power. However, results from Table 24 weaken the conclusion that hypothesis 9 can be accepted. For example, average variability in operating margins within firms of a given industry loses significant explanatory power (t-stat=1.62) when concentration of wealth creation is based on inflation instead of the risk-free asset. Thus, results from robustness checks complicate drawing a consistent conclusion on whether to accept or reject hypothesis 9.

Figure 20: MEP interest coverage

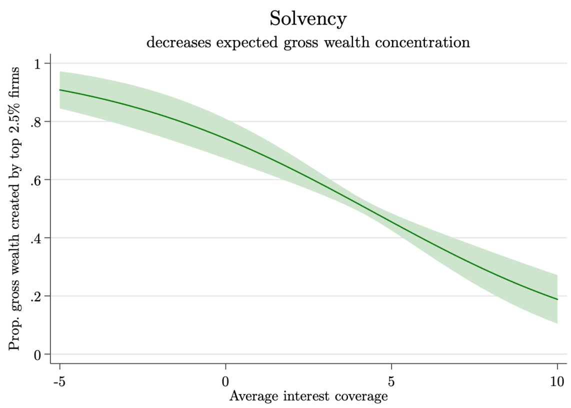



Figure 21: MEP ROCE

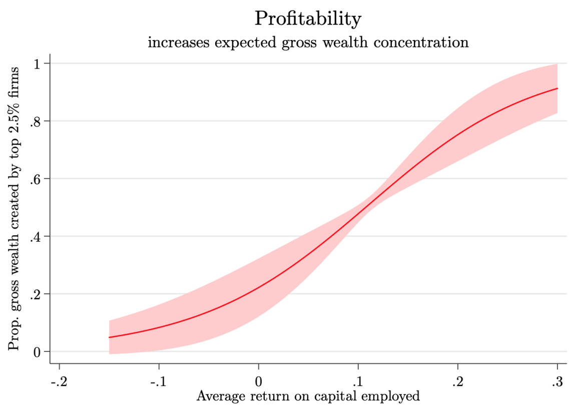


Table 22: Cross-industry outperformance of CPI

| | Proportion of firms outperforming CPI over their lifetime | | | | |
|---|---|---|---|---|---|
| | 1926-2018 | 1970-2018 | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Standard deviation | $-11.53^{***}$ | $-9.557^{***}$ | | | $-14.55^{***}$ |
| | $(-8.49)$ | $(-6.32)$ | | | $(-9.42)$ |
| Skewness | 0.187 | $-0.290$ | | | 0.221 |
| | (0.62) | $(-0.91)$ | | | (1.48) |
| Mean return | $23.52^{***}$ | $23.08^{***}$ | | | $25.71^{***}$ |
| | (3.90) | (4.31) | | | (5.32) |
| Capital ratio | | | $2.032^{*}$ | | $1.119^{*}$ |
| | | | (2.32) | | (2.46) |
| Asset turnover | | | $-0.258$ | | $0.319^{**}$ |
| | | | $(-1.43)$ | | (2.94) |
| Quick ratio | | | 0.143 | | 0.115 |
| | | | (0.54) | | (1.11) |
| ROCE | | | 2.629 | | $-0.534$ |
| | | | (0.71) | | $(-0.43)$ |
| Operating margin | | | 0.703 | | $1.953^{***}$ |
| | | | (0.72) | | (3.78) |
| Cash flow/debt | | | $-2.778^{**}$ | | 0.722 |
| | | | $(-3.13)$ | | (1.45) |
| Interest coverage | | | 0.0903 | | 0.00302 |
| | | | (1.75) | | (0.15) |
| Capital ratio SD | | | | 1.529 | $-1.821$ |
| | | | | (0.39) | $(-1.02)$ |
| Asset turnover SD | | | | $-2.238$ | $-1.268$ |
| | | | | $(-1.70)$ | $(-1.47)$ |
| Quick ratio SD | | | | $-0.432$ | $-0.0203$ |
| | | | | $(-1.54)$ | $(-0.13)$ |
| ROCE SD | | | | $-4.426$ | $7.542^{***}$ |
| | | | | $(-1.56)$ | (4.81) |
| Operating margin SD | | | | 0.749 | $0.531^{**}$ |
| | | | | (1.94) | (3.17) |
| Cash flow/debt SD | | | | $-3.954^{**}$ | $-2.072^{***}$ |
| | | | | $(-3.14)$ | $(-3.85)$ |
| Interest coverage SD | | | | $0.0231^{**}$ | 0.000129 |
| | | | | (3.03) | (0.03) |
| Constant | $1.652^{***}$ | $1.652^{***}$ | $-1.118$ | $1.075^{*}$ | $0.956^{**}$ |
| | (7.40) | (8.27) | $(-1.50)$ | (2.06) | (2.83) |
| Observations | 50 | 50 | 49 | 50 | 49 |

$t$ statistics in parentheses
This fractional logistic regression measures the effect of the short-term mean, standard deviation (SD) and skewsness of stock returns on long-term outperformance over CPI of individual stocks across industries. The mean, SD and skewness are calculated for each stock and averaged to get to the industry figures. In regressions (3), (4) and (5) corporate performance figures and the standard deviation of those figures is also taken into account.
$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Table 23: Cross-industry gross wealth concentration by top 2.5% of firms, based on CPI

| | Proportion of gross wealth created by top 2.5% of firms | | | | |
|---|---|---|---|---|---|
| | 1926-2018 | 1970-2018 | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Standard deviation | 7.831* | 8.884 | | | 4.830 |
| | (2.06) | (1.80) | | | (1.08) |
| Skewness | −1.338 | −1.199 | | | −0.687 |
| | (−1.74) | (−1.28) | | | (−1.16) |
| Mean return | 18.24 | 17.05 | | | 51.30*** |
| | (1.25) | (1.09) | | | (3.42) |
| Capital ratio | | | −4.424*** | | −6.027*** |
| | | | (−3.46) | | (−5.04) |
| Asset turnover | | | −0.112 | | −0.274 |
| | | | (−0.37) | | (−0.74) |
| Quick ratio | | | −0.209 | | −0.888** |
| | | | (−0.62) | | (−2.77) |
| ROCE | | | 13.13*** | | 13.15*** |
| | | | (4.17) | | (4.00) |
| Operating margin | | | 0.339 | | −1.918 |
| | | | (0.20) | | (−0.91) |
| Cash flow/debt | | | −2.810 | | −1.594 |
| | | | (−1.44) | | (−1.02) |
| Interest coverage | | | −0.208** | | −0.283*** |
| | | | (−2.73) | | (−5.07) |
| Capital ratio SD | | | | 1.886 | 3.932 |
| | | | | (0.30) | (0.79) |
| Asset turnover SD | | | | −3.713 | −6.160* |
| | | | | (−1.74) | (−2.17) |
| Quick ratio SD | | | | 0.213 | 0.415 |
| | | | | (0.42) | (0.84) |
| ROCE SD | | | | 7.871* | 12.99** |
| | | | | (2.13) | (2.70) |
| Operating margin SD | | | | −1.790** | −2.330*** |
| | | | | (−3.27) | (−4.05) |
| Cash flow/debt SD | | | | 2.474 | 2.934 |
| | | | | (1.32) | (1.79) |
| Interest coverage SD | | | | 0.00124 | −0.0306* |
| | | | | (0.10) | (−2.04) |
| Constant | −0.105 | −0.537 | 1.357 | −0.497 | 2.691** |
| | (−0.17) | (−0.94) | (1.36) | (−0.70) | (2.86) |
| Observations | 50 | 50 | 49 | 50 | 49 |

*t* statistics in parentheses
This fractional logistic regression measures the effect of the short-term mean, standard deviation (SD) and skewness of stock returns on long-term concentration of gross wealth creation across industries. The mean, SD and skewness are calculated for each stock and averaged to get to the industry figures. In regressions (3), (4) and (5) corporate performance figures and the standard deviation of those figures is also taken into account.
$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Table 24: Cross-industry net wealth concentration, based on CPI

| | Proportion of firms responsible for 100% of net wealth creation | | | | |
|---|---|---|---|---|---|
| | 1926-2018 | 1970-2018 | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Standard deviation | $-20.20^{***}$ | $-12.29$ | | | $-46.39^{***}$ |
| | $(-3.80)$ | $(-1.66)$ | | | $(-9.62)$ |
| Skewness | 0.931 | $-0.941$ | | | 0.183 |
| | (0.95) | $(-0.70)$ | | | (0.31) |
| Mean return | $58.60^{***}$ | $63.82^{**}$ | | | 24.58 |
| | (4.42) | (2.82) | | | (1.14) |
| Capital ratio | | | 2.957 | | 2.246 |
| | | | (1.13) | | (1.22) |
| Asset turnover | | | $-0.963^{*}$ | | 0.145 |
| | | | $(-2.42)$ | | (0.27) |
| Quick ratio | | | $-0.729$ | | $-0.458$ |
| | | | $(-0.97)$ | | $(-0.94)$ |
| ROCE | | | $15.12^{*}$ | | 4.487 |
| | | | (2.53) | | (0.80) |
| Operating margin | | | $-5.668^{*}$ | | 0.594 |
| | | | $(-1.99)$ | | (0.26) |
| Cash flow/debt | | | $-3.369$ | | 0.991 |
| | | | $(-1.23)$ | | (0.46) |
| Interest coverage | | | 0.0573 | | 0.00500 |
| | | | (0.38) | | (0.04) |
| Capital ratio SD | | | | 23.62 | $17.43^{*}$ |
| | | | | (1.42) | (2.55) |
| Asset turnover SD | | | | $-3.949$ | 0.627 |
| | | | | $(-0.99)$ | (0.19) |
| Quick ratio SD | | | | $-3.978^{**}$ | $-1.046$ |
| | | | | $(-2.92)$ | $(-1.22)$ |
| ROCE SD | | | | $-24.42^{**}$ | 7.003 |
| | | | | $(-2.91)$ | (1.16) |
| Operating margin SD | | | | $3.391^{*}$ | 1.055 |
| | | | | (2.31) | (1.62) |
| Cash flow/debt SD | | | | $-0.110$ | $6.548^{**}$ |
| | | | | $(-0.03)$ | (3.14) |
| Interest coverage SD | | | | $0.0650^{**}$ | 0.0119 |
| | | | | (3.10) | (0.67) |
| Constant | $-0.197$ | 0.00482 | $-2.772$ | $-1.116$ | $-0.143$ |
| | $(-0.26)$ | (0.01) | $(-1.36)$ | $(-0.63)$ | $(-0.12)$ |
| Observations | 50 | 50 | 49 | 50 | 49 |

*t* statistics in parentheses
This fractional logistic regression measures the effect of the short-term mean, standard deviation (SD) and skewsness of stock returns on long-term concentration of net wealth creation across industries. The mean, SD and skewness are calculated for each stock and averaged to get to the industry figures. In regressions (3), (4) and (5) corporate performance figures and the standard deviation of those figures is also taken into account.
$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

# 8 Limitations and opportunities

This paper finds evidence that concentration of long-term wealth creation across US industries is positively associated with short-term volatility and is negatively associated with mean returns. These results are in line with the theory from Bessembinder (2018) and Farago and Hjalmarsson (2019) that short-term volatility induces skewness in long-term stock returns which implies that a few stocks experience exceptional returns while the majority have below-average returns. Some evidence is found that better corporate performance of firms in a given industry is associated with a lower concentration in wealth creation. This is in line with the theory that industries where firms have better corporate performance have a larger proportion of firms that outperform the risk-free asset and thus create wealth which reduces concentration of wealth. Authors such as Fama and French (2015), Novy-Marx (2013) and Piotroski (2000) also found that portfolios of stock with better corporate performance characteristic outperform portfolios with worse characteristics. The exact relation between variability of corporate performance within firms of a given industry and concentration of wealth creation is not unanimously positive or negative. Results differ based on what financial ratio is used. However, we do accept the hypothesis that variability in corporate performance can explain cross-industry variation in concentration of wealth creation.

The results of this research are subject to several limitations and can potentially benefit from improvements. This section elaborates on those limitations and highlights opportunities for further research into the topic of long-term stock returns and concentration of wealth creation.

First, categorizing firms into industries is prone to errors. For instance, this paper uses Standard Industry Classification (SIC) codes which were officially replaced by North American Industry Classification System (NAICS) codes in 1997[16]. Therefore, definitions to classify a business to a certain SIC code could be outdated. However, this problem is partly mitigated, because the Securities and Exchange Commission (SEC) still uses SIC codes and requires them in fillings of listed companies[17]. Compustat scrapes SIC codes from annual and quarterly reports filled at the SEC, which makes it a reasonably reliable source for SIC codes. CRSP relies on Interactive Data Corporation, a company owned by ICE, the owner of the New York Stock Exchange. This paper deems that a reliable source for SIC codes. Furthermore, one could bring up arguments against grouping firms into industries, because it is based on the assumption they operate in the same market and compete with each other. However, firms might not operate in the same market. For instance, some firms are active overseas while others focus on their domestic market. Additionally, firms could produce different products while still being in the same industry, especially if they are in a niche. Lastly, some firms are hard to classify, because they have different business segments. For example, Johnson & Johnson is active in pharmaceutical products, medical equipment and consumer healthcare but is classified as being in the pharmaceutical industry. An opportunity for future research is to use a more granular or detailed industry classification system to check the robustness of the results of the sample applied in this research.

[16]Source: https://www.census.gov/eos/www/naics/
[17]Source: https://www.sec.gov/info/edgar/siccodes.htm

Second, this paper may suffer from omitted variable bias, because competition between firms and pricing power of firms is not taken into account. Grullon et al. (2019) measure industry concentration levels by using the Herfindahl-Hirschman index which measures the degree of competition in product markets. They also assess pricing power by looking at expanding operating margins. Grullon et al. (2019) find that industry concentration has increased significantly over the last two decades, market concentration increases profit margins and firms in concentrated industries outperform firms in less concentrated industries. On the contrary, Hou and Robinson (2006) utilize the Herfindahl-Hirschman index as well and find that firms in more concentrated industries earn experience lower stock returns. The authors do find evidence that firms in concentrated industries display abnormal profitability. Hence, existing literature relating product market competition to stock returns is contradictory. This paper decided against using the Herfindahl-Hirschman index. The reason is that Ali et al. (2008) find that industry concentration calculations using Compustat data, which only contains firms listed on stock exchange, are unreliable. Furthermore, Ali et al. (2008) replicate several studies, including Hou and Robinson (2006). Ali et al. (2008) find that results become insignificant when using more reliable US Census data, because of the omission of private firms and the low correlation between Compustat and US Census data. Future research could include industry concentration metrics by using US Census data. However, applying Census data to an industry classification such as the one in this paper requires extremely extensive hand collection of data.

Third, this research omits privately owned firms, because price, shares outstanding and dividend data is unavailable or infrequent for private firms. Additionally, private firms rarely publish accounting statements. Hence, corporate performance figures cannot be calculated for private firms. Industries with large successful private companies could be less concentrated than the figures calculated in this paper. For instance, Mars, the 6th largest privately held company in the US and maker of confectionery and pet food, would be included in the food products industry[18]. Another example is Cargill, the largest privately owned company in the United States according to Forbes. With the inclusion of Cargill, calculations for the agriculture industry would certainly be different. Opportunities to include privately owned firms are limited, because valuation is subjective and infrequent.

Last, several other opportunities to increase the understanding of long-term stock behaviour arise. First, this research could be extended to more countries or to the global stock market. Second, other ways of measuring wealth creation or returns matching the experience of investors can be utilized. For example, an internal rate of return (IRR) could be calculated for every stock as if investors invest at IPO and sell at the end of the sample. Third, recently published research from Bessembinder (2020) finds that concentration in wealth concentration has been increasing in recent decades. Research into why that is the case could be insightful. Could determinants of concentration in wealth creation found in this paper play a role in that development? A new approach to the puzzle of concentration in wealth creation could be to take a time-series approach by subdividing the sample into shorter horizons. It would increase the sample size and allow analysis of why the degree of concentration in wealth concentration changes over time.

[18]Source: https://www.forbes.com/largest-private-companies/list/

# 9 Conclusion

A large body of literature has sought to explain equity returns by focusing on relatively short monthly and annual horizons. However, one could question whether such a short horizon matches with the investment horizon of the typical investor. Both institutional and retail investors are likely to have much longer investment horizons, spanning multiple years if not decades. Recent research by Bessembinder (2018) shows that long-term individual stock returns are asymmetric and the average return of the stock market in aggregate obscures the extremely skewed underlying distribution of long-term stock returns. Most stocks are unable to outperform the risk-free rate. Additionally, extreme skewness induces concentration of wealth creation[19] in a few stocks.

Severe asymmetry and concentration in wealth creation is risky for investors holding a concentrated portfolio of stocks, because they are likely to miss long-term winners. As a matter of fact, concentration in wealth creation increases the probability of investors holding a portfolio other than the market portfolio to underperform the market. Therefore, it is beneficial for practitioners to know what determines concentration of wealth creation. This paper takes a novel approach by investigating concentration, skewness and stock's inability to outperform the risk-free rate in US industries. It is the first to investigate long-term concentration of wealth creation in US industries. Both institutional and retail practitioners can make more informed choices if they know which industries and markets are especially difficult for stock selection, i.e. suffer from severe asymmetry in long-term returns. Hence, this papers' research question is:

**Do US industries suffer from concentration in wealth creation and what are the empirical determinants of cross-industry variation in concentration of wealth creation?**

By using US stock returns from CRSP in the period from 1926 to 2018 combined with industry categorization this paper is able to show that long-term skewness and concentration of wealth creation differ widely across industries. Skewness in buy-and-hold stock returns generally increases with longer horizons. At a monthly horizon, skewness is roughly 6.5, increasing to 20 on an annual horizon, 30 on a decade horizon and 109 at the lifetime horizon. Industries suffering from extreme positive skewness in lifetime buy-and-hold returns are the banking, business services, trading, wholesale and retail industry. Logically, large differences exist between mean and median buy-and-hold returns, in 23 out of 50 industries median lifetime buy-and-hold returns are negative whereas in all industries mean lifetime returns are positive. We accept hypothesis 1 and conclude that long-term returns across most US industries are extremely positively skewed.

It should come as no surprise that if the median return of an industry is close to zero or negative a majority of stocks underperform the risk-free rate. Hypothesis 2 claims that the majority of stocks across US industries are unable to achieve a higher buy-and-hold return over their lifetime than the return of the risk-free asset. The ability of stocks from a given industry to outperform the risk-free rate or inflation over their lifetime varies significantly, ranging from an abysmal

[19] To recap, wealth creation by a stock is defined as: "the difference between the present US dollar value of investors' actual investments in the stock and the value that would have been obtained if the same capital investments had earned one-month US Treasury Bill returns."

22% to an acceptable 74%. In the precious metals, recreation and computer industry less than 30% of stocks were able to outperform the risk-free rate over their lifetime. On the contrary, in the utilities, tobacco products and candy & soda industry more than 65% of stocks were able to outperform the risk-free rate over their lifetime. This paper fails to reject hypothesis 2 in a t-test. Thus, the majority of stocks in a given US industry destroy wealth over their lifetime.

The second part of the research question is indirectly answered by assessing cross-industry variation in the proportion of stocks that are able to outperform the risk-free rate over their lifetime. Hypotheses 3, 4 and 5 are summarized by asserting that short-term return characteristics (3), corporate performance of industries (4) and variability in corporate performance within firms (5) can explain cross-industry differences in the proportion of stocks that are able to outperform the risk-free asset over their lifetime. Thus, we hypothesize that these three sets of empirical determinants can explain why some industries contain a larger proportion of stocks that create wealth over their lifetime than other industries.

Regarding hypothesis 3, Bessembinder et al. (2019) and Farago and Hjalmarsson (2019) explain that short-term return volatility and short-term return skewness are drivers for long-term stock return skewness. Hence, we expect both to be negatively associated with the proportion of stocks that outperform the risk-free rate over their lifetime. Additionally, mean returns are included as a control variable since it is intuitive that higher short-term return result in higher long-term returns, i.e. a higher proportion of stocks that outperform the risk-free rate. We use fractional outcome regressions to test hypotheses where the dependent variable is a proportion, because proportions have a range of [0,1] . Our regressions are non-linear and interpretation of coefficients is complex. Therefore, we use marginal effects plots to visualize the relation between the dependent variable and independent variables.

In our regression analysis, we find that short-term return volatility in a given industry is negatively associated with the proportion of firms that outperform the risk-free rate over their lifetime. Additionally, higher short-term mean returns in a given industry are associated with a larger proportion of stocks outperforming the risk-free rate over their lifetime. Therefore, investors should be aware that industries with high short-term volatility and low short-term mean returns are expected to contain a large proportion of stocks that are unable to outperform the risk-free rate. Hence, the risk investors assume by investing in these industries is unlikely to be rewarded with a return above the risk-free asset. Examples of those industries are the entertainment, pharmaceutical and computer industry.

With respect to hypothesis 4, we use capitalization, solvency, liquidity, efficiency and profitability to assess whether corporate performance of firms in a given industry can explain variation in the proportion of stocks that are able to outperform the risk-free rate over their lifetime. We expect that better corporate performance is positively related to the proportion of stocks that outperform the risk-free rate, because they have a lower probability of bankruptcy (Beaver (1966); Altman (1968)) and their more profitable and stable businesses are expected to be rewarded with higher stock returns (Fama and French (2015); Novy-Marx (2013); George and Hwang (2010)). Our empirical results show that the relationship between corporate performance in industries and the proportion of stocks that outperform the risk-free rate is not unanimously negative or positive. Industries with more efficient, profitable and leveraged firms are expected to have a higher proportion of stocks that outperform the risk-free rate over their lifetime. The tobacco industry is the only industry fulfilling all those criteria. These results learn investors they should be careful when selecting stocks in industries with mediocre corporate performance as a large proportion of stocks in those industries is expected to underperform the risk-free rate. Empirical determinants based on corporate performance can partly explain cross-industry variation in the proportion of stocks that outperform the risk-free rate. Therefore, we accept hypothesis 4 with discretion.

With respect to hypothesis 5, this paper expects that variability of corporate performance within firms of a given industry is negatively associated with the proportion of stocks that is able to outperform the risk-free rate. More variability in corporate performance could result in higher default rates of firms (Dambolena and Khoury (1980);Betts and Belhoul (1987)). Unfortunately, no existing research exists that relates industry competition to variability in corporate performance. However, common sense would dictate that increased variability in corporate performance could theoretically be an indicator of increased competition. Since we only cover the public market, this paper decided against using industry competition metrics as they are likely to suffer from omission of private firms (Ali et al. (2008)).

Empirical results from testing hypothesis 5 show that variability of profitability in a given industry increases the expected proportion of firms that outperform the risk-free rate. A potential explanation could be found in results from Mauboussin and Callahan (2013) who find that firms with rapidly increasing profitability are rewarded in the stock market. In other words, their shift from a value-destroying business to a value-creating business generates substantial returns for shareholders (Mauboussin and Callahan (2013)). Additionally, we find that variability in solvency within firms of a given industry is associated with a lower proportion of firms that outperform the risk-free rate which is probably caused by a larger rate of default. Therefore, investor should be cautious in industries where firms have a high variability in solvency such as the pharmaceutical, medical equipment and precious metals industry. Furthermore, we accept hypothesis 5.

Results of testing hypotheses 1-5 investors learn investors that they are likely to face difficulties in stock selection in most US industries due to extreme skewness and the fact that the majority of stocks underperform the risk-free rate. They can already improve their chances of success by avoiding industries with high short-term return volatility, return skewness and by choosing to invest in industries with good corporate performance. Results from testing hypothesis 6 answer the first part of the research question directly by showing that most US industries suffer from concentration in wealth creation.

Hypothesis 6 states that a high degree of concentration in wealth creation is present across US industries. Evidence in support for this hypothesis is found by the fact that in the most concentrated industries less than 1% of firms is responsible for all net wealth creation. In many industries just a handful of stocks is responsible for the vast majority of wealth creation. Extreme concentration of wealth creation is present in the electronic equipment, computer software and computer industry. In less concentrated industries up to 43% of firms is responsible for all net wealth creation, wealth creation was least concentrated in the tobacco, utilities and candy & soda industry. Furthermore, this paper finds that a number of industries have not been able to create any wealth in the entire period from 1926 to 2018, such as the fabricated products, coal and precious metals industry. Testing hypothesis 6 formally in a t-test shows that it is rejected if a stringent threshold is used, i.e. $< 5\%$ of firms is responsible for all net wealth creation across US industries. This result is heavily influenced by a few industries where concentration of wealth creation was relatively low. Therefore, we do accept hypothesis 6 as the majority of industries suffer from extreme concentration in wealth creation.

Hypotheses 7, 8 and 9 use the same empirical determinants as hypotheses 3, 4 and 5 to investigate whether short-term return characteristics, corporate performance and variability in corporate performance can explain cross-industry variation in concentration of wealth creation. Hence, hypotheses 7, 8 and 9 answer the second part of the research question directly. We generally expect opposite relations compared hypotheses 3-5. Because if a larger proportion of firms is able to outperform the risk-free rate over their lifetime, concentration in wealth creation is likely to be lower.

Results with respect to hypothesis 7 indicate that higher short-term volatility and lower short-term mean returns in a given industry leads to more concentration of wealth creation. Therefore, industries with these characteristics such as the entertainment, pharmaceutical and computer industry are prone to concentration in wealth creation; increasing the difficulty of stock selection. Thus, quite intuitively, long-term investors such as pension funds who think about investing in these industries should possess substantial stock picking skills to select long-term winners as there are few. We have tested hypothesis 7 again in the robustness section and evidence was mixed. However, we deem the evidence strong enough to accept it, because results with respect to concentration in net wealth creation were convincing.

Regarding hypothesis 8, empirical results indicate that better corporate performance of firms in a given industry does not necessarily imply a lower degree of wealth concentration. This paper performs four regression to test the relationship between corporate performance and concentration of wealth creation in US industries. With respect to concentration in gross wealth creation we find evidence that more leverage, higher liquidity and better solvency decrease concentration in gross wealth creation. Industries where firms are more leveraged could be more regulated, such as the banking, utilities and tobacco industry (our data shows that these industries have high leverage and low concentration in wealth creation) which would induce a more even distribution of wealth creation. Altman (1968), Beaver (1966) and Piotroski (2000) show that firms with higher liquidity and better solvency tend to have a lower default rate and experience higher stock returns as a group which could lower concentration in wealth creation. However, variables in other regressions that relate concentration in net wealth creation to corporate performance are largely insignificant. Thus, we are inclined to reject hypothesis 8.

Lastly, we find mixed results when testing hypothesis 9 which states that variability in corporate performance can explain cross-industry variation in concentration of wealth creation. Theoretically, a higher variability in corporate performance such as profitability could point to a more competitive industry where industry concentration is lower (Gaspar and Massa (2006)). However, variability in corporate performance can also reflect a higher probability of bankruptcy (Dambolena and Khoury (1980)) which would increase concentration in wealth creation to a few surviving firms. The majority of results in this paper point to the first theory. More variability in leverage, profitability and solvency in a given industry is related to lower concentration of wealth creation. However, industries with more variability in corporate performance might experience less concentration in wealth creation they are also associated with having a lower proportion of stocks that outperform the risk-free rate. This results is logical if the firms that do outperform the risk-free rate have a relatively even distribution of wealth creation. Hence, investors should still be cautious when investing in industries with high variability in corporate performance.

To sum up, the empirical evidence in this paper sheds light on the puzzle of long-term stock return skewness, concentration of wealth concentration and the inability of most stocks to outperform the risk-free rate and inflation. This paper is the first to assess concentration in wealth creation in US industries. It is also the first to show that corporate performance has explanatory power for the cross-section of concentration in wealth creation across US industries. The answer to the research question is that most US industries suffer from concentration in wealth creation. Empirical determinants of cross-industry variation in concentration of wealth creation are found in three categories: short-term return characteristics, corporate performance and variability in corporate performance. Investors should realize that in most industries just a few winners are responsible for the vast majority of wealth creation. This result implies that the difficulty of stock selection fluctuates between industries. They now have some tools to determine which industries are likely to suffer from concentration in wealth creation.

## References

Alfen, T. V. (2017). SICFF: Stata module to create Fama French Industry Variable from SIC Code. Statistical Software Components, Boston College Department of Economics.

Ali, A., Klasa, S., and Yeung, E. (2008). The limitations of industry concentration measures constructed with compustat data: Implications for finance research. *The Review of Financial Studies*, 22(10):3839–3871.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4):589–609.

Arditti, F. D. and Levy, H. (1975). Portfolio efficiency analysis in three moments: The multiperiod case. *The Journal of Finance*, 30(3):797–809.

Baum, C. F. (2008). Stata tip 63: Modelling proportions. *The Stata Journal*, 8(2):299–303.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of accounting research*, pages 71–111.

Bessembinder, H. (2018). Do stocks outperform treasury bills? *Journal of financial economics*, 129(3):440–457.

Bessembinder, H. (2020). Wealth creation in the us public stock markets 1926 to 2019. *Available at SSRN*.

Bessembinder, H., Chen, T.-F., Choi, G., and Wei, K. (2019). Do global stocks outperform us treasury bills? *Te-Feng and Choi, Goeun and Wei, Kuo-Chiang (John), Do Global Stocks Outperform US Treasury Bills*.

Betts, J. and Belhoul, D. (1987). The effectiveness of incorporating stability measures in company failure models. *Journal of Business Finance & Accounting*, 14(3):323–334.

Blitz, D. (2014). The dark side of passive investing. *Journal of Portfolio Management*, 41(1):1–4.

Cameron, A. C. and Trivedi, P. K. (2010). Microeconometrics using stata (revised ed.). *Number musr in Stata Press books. StataCorp LP*.

Conrad, J., Dittmar, R. F., and Ghysels, E. (2013). Ex ante skewness and expected stock returns. *The Journal of Finance*, 68(1):85–124.

Dambolena, I. G. and Khoury, S. J. (1980). Ratio stability and corporate failure. *The Journal of Finance*, 35(4):1017–1026.

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

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds.

*Journal of.*

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

Fama, E. F. and French, K. R. (2018). Long-horizon returns. *The Review of Asset Pricing Studies*, 8(2):232–252.

Farago, A. and Hjalmarsson, E. (2019). Compound returns. *Available at SSRN 3398501.*

French, K. R. (2008). Presidential address: The cost of active investing. *The Journal of Finance*, 63(4):1537–1573.

Friend, I. and Blume, M. (1970). Measurement of portfolio performance under uncertainty. *The American economic review*, 60(4):561–575.

Gaspar, J.-M. and Massa, M. (2006). Idiosyncratic volatility and product market competition. *The Journal of Business*, 79(6):3125–3152.

George, T. J. and Hwang, C.-Y. (2010). A resolution of the distress risk and leverage puzzles in the cross section of stock returns. *Journal of Financial Economics*, 96(1):56–79.

Gharbi, S., Sahut, J.-M., and Teulon, F. (2014). R&d investments and high-tech firms' stock return volatility. *Technological Forecasting and Social Change*, 88:306–312.

Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo maximum likelihood methods: Theory. *Econometrica: Journal of the Econometric Society*, pages 681–700.

Grullon, G., Larkin, Y., and Michaely, R. (2019). Are us industries becoming more concentrated? *Review of Finance*, 23(4):697–743.

Haugen, R. A., Baker, N. L., et al. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, 41(3):401–439.

Heaton, J., Polson, N., and Witte, J. H. (2017). Why indexing works. *Applied Stochastic Models in Business and Industry*, 33(6):690–693.

Hou, K. and Robinson, D. T. (2006). Industry concentration and average stock returns. *The Journal of Finance*, 61(4):1927–1956.

Hou, K., Xue, C., and Zhang, L. (2017). Replicating anomalies. *The Review of Financial Studies.*

Ikenberry, D., Shockley, R., and Womack, K. (1998). Why active fund managers often underperform the s&p 500: The impact of size and skewness. *Journal of Private Portfolio Management*, 1(1):13–26.

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

Kraus, A. and Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *The Journal of finance*, 31(4):1085–1100.

Mauboussin, M. and Callahan, D. (2013). Economic returns, reversion to the mean and total shareholder returns. *Global Financial Strategies*, 2013:1–14.

Mazzucato, M. and Tancioni, M. (2012). R&d, patents and stock return volatility. *Journal of Evolutionary Economics*, 22(4):811–832.

McCullagh, P. and Nelder, J. (1989). Generalized linear models new york chapman & hall. *McCullagh2Generalized Linear Models1989*.

Mehra, R. and Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of monetary Economics*, 15(2):145–161.

Morawski, J., Rehkugler, H., and Füss, R. (2008). The nature of listed real estate companies: property or equity market? *Financial Markets and Portfolio Management*, 22(2):101.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology*, 47(1):90–100.

Neuberger, A. and Payne, R. (2018). The skewness of the stock market at long horizons. *Available at SSRN 3173581*.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, pages 109–131.

Papke, L. E. and Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of applied econometrics*, 11(6):619–632.

Pástor, L. and Pietro, V. (2003). Stock valuation and learning about profitability. *The Journal of Finance*, 58(5):1749–1789.

Piotroski, J. D. (2000). Value investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*, pages 1–41.

Porter, M. E. (1980). Industry structure and competitive strategy: Keys to profitability. *Financial analysts journal*, 36(4):30–41.

Royston, P. (2013). marginscontplot: Plotting the marginal effects of continuous predictors. *The Stata Journal*, 13(3):510–527.

Simkowitz, M. A. and Beedles, W. L. (1978). Diversification in a three-moment world. *Journal* *of Financial and Quantitative Analysis*, 13(5):927–941.

Soliman, M. T. (2008). The use of dupont analysis by market participants. *The Accounting Review*, 83(3):823–853.

StataCorp (2017). *Stata Statistical Software: Release 15*. StataCorp LLC, College Station, TX.

Vail, S. V. and Valavanis, S. (1959). *Econometrics: An Introduction to Maximum Likelihood Methods*. McGraw-Hill.

Wedderburn, R. W. (1974). Quasi-likelihood functions, generalized linear models, and the gauss—newton method. *Biometrika*, 61(3):439–447.

Williams, R. (2012). Using the margins command to estimate and interpret adjusted predictions and marginal effects. *Stata Journal*, 12(2):308–331(24).

# Appendix A: buy-and-hold return summary statistics

Table 25: CRSP stocks buy-and-hold returns at monthly horizon grouped by industry during 1970-2018

| | Monthly buy-and-hold returns | | | | |
|---|---|---|---|---|---|
| | Mean | Median | Standard deviation | Skewness | Obs |
| **Fama-French 49 Industries** | | | | | |
| Agriculture | 0.0059 | 0.0000 | 0.186 | 3.549 | 11,325 |
| Food Products | 0.0114 | 0.0000 | 0.142 | 3.325 | 49,983 |
| Candy & Soda | 0.0126 | 0.0051 | 0.122 | 2.420 | 9,435 |
| Beer & Liquor | 0.0090 | 0.0000 | 0.142 | 2.414 | 11,773 |
| Tobacco Products | 0.0145 | 0.0060 | 0.140 | 4.096 | 4,645 |
| Recreation | 0.0058 | 0.0000 | 0.191 | 3.614 | 29,229 |
| Entertainment | 0.0083 | 0.0000 | 0.224 | 7.226 | 41,040 |
| Printing & Publishing | 0.0109 | 0.0000 | 0.157 | 4.885 | 32,377 |
| Consumer Goods | 0.0097 | 0.0000 | 0.162 | 5.953 | 62,715 |
| Apparel | 0.0078 | 0.0000 | 0.164 | 2.634 | 41,094 |
| Healthcare | 0.0119 | 0.0000 | 0.204 | 4.143 | 54,235 |
| Medical Equipment | 0.0104 | 0.0000 | 0.223 | 18.706 | 80,300 |
| Pharmaceutical Products | 0.0122 | 0.0000 | 0.240 | 7.408 | 126,592 |
| Chemicals | 0.0115 | 0.0000 | 0.153 | 2.555 | 54,015 |
| Rubber and Plastic Products | 0.0109 | 0.0000 | 0.169 | 2.134 | 25,845 |
| Textiles | 0.0075 | 0.0000 | 0.157 | 1.919 | 25,241 |
| Construction Materials | 0.0118 | 0.0000 | 0.159 | 3.683 | 80,836 |
| Construction | 0.0086 | 0.0000 | 0.192 | 5.066 | 36,598 |
| Steel Works | 0.0101 | 0.0000 | 0.176 | 29.886 | 44,138 |
| Fabricated Products | 0.0105 | 0.0000 | 0.169 | 5.009 | 12,791 |
| Machinery | 0.0118 | 0.0000 | 0.167 | 4.868 | 100,449 |
| Electrical Equipment | 0.0107 | 0.0000 | 0.221 | 5.788 | 72,053 |
| Automobiles and Trucks | 0.0080 | 0.0000 | 0.149 | 1.415 | 43,569 |
| Aircraft | 0.0135 | 0.0000 | 0.150 | 2.875 | 16,152 |
| Shipbuilding, Railroad Equipment | 0.0115 | 0.0000 | 0.151 | 1.675 | 5,473 |
| Defense | 0.0142 | 0.0036 | 0.154 | 1.464 | 5,248 |
| Precious Metals | 0.0067 | 0.0000 | 0.220 | 2.973 | 35,024 |
| Mines | 0.0053 | 0.0000 | 0.201 | 5.941 | 25,599 |
| Coal | 0.0072 | 0.0000 | 0.195 | 3.498 | 5,715 |
| Petroleum and Natural Gas | 0.0073 | 0.0000 | 0.204 | 8.303 | 151,850 |
| Utilities | 0.0108 | 0.0088 | 0.085 | 3.361 | 95,553 |
| Communication | 0.0099 | 0.0000 | 0.219 | 5.887 | 75,464 |
| Personal Services | 0.0083 | 0.0000 | 0.201 | 8.123 | 32,499 |
| Business Services | 0.0104 | 0.0000 | 0.204 | 4.067 | 178,404 |
| Computers | 0.0110 | 0.0000 | 0.221 | 3.488 | 82,629 |
| Computer Software | 0.0117 | 0.0000 | 0.252 | 5.144 | 143,090 |
| Electronic Equipment | 0.0131 | 0.0000 | 0.206 | 5.138 | 151,875 |
| Measuring and Control Equipment | 0.0139 | 0.0000 | 0.202 | 9.927 | 60,370 |
| Business Supplies | 0.0109 | 0.0000 | 0.147 | 7.787 | 30,702 |
| Shipping Containers | 0.0122 | 0.0000 | 0.146 | 5.893 | 15,973 |
| Transportation | 0.0085 | 0.0000 | 0.169 | 3.503 | 74,431 |
| Wholesale | 0.0094 | 0.0000 | 0.194 | 5.067 | 120,724 |
| Retail | 0.0091 | 0.0000 | 0.174 | 4.417 | 154,066 |
| Restaurant, Hotels, Motels | 0.0079 | 0.0000 | 0.167 | 3.420 | 65,252 |
| Banking | 0.0096 | 0.0034 | 0.127 | 4.338 | 241,656 |
| Insurance | 0.0110 | 0.0025 | 0.129 | 4.392 | 92,184 |
| Real Estate | 0.0075 | 0.0000 | 0.142 | 7.446 | 90,598 |
| Trading | 0.0124 | 0.0000 | 0.136 | 4.030 | 221,326 |
| Other | 0.0087 | 0.0000 | 0.201 | 3.672 | 11,974 |
| Not available | 0.0034 | 0.0000 | 0.202 | 2.767 | 17,667 |
| Total | 0.0103 | 0.0000 | 0.183 | 6.519 | 3,251,776 |

All common stocks and real estate investment trusts (REITs), occurring at some point in time, in CRSP from January 1970 until December 2018 are included. Stocks are assigned to an industry based on their Standard Industry Classification (SIC) code. Kenneth French's data library is used to structure which SIC codes belong to which industry. Companies are allowed to switch industries multiple times. If companies switch back to an industry they were previously in, this is treated as a separate run of returns to prevent faulty linking of returns. Mean and median returns depicted in the table are monthly buy-and-hold returns including dividends and delisting returns.

Table 26: CRSP stocks buy-and-hold returns at annual horizon grouped by industry during 1970-2018

| | Annual buy-and-hold returns | | | | |
|---|---|---|---|---|---|
| | Mean | Median | Standard deviation | Skewness | Obs |
| **Fama-French 49 Industries** | | | | | |
| Agriculture | 0.0649 | 0.0000 | 0.638 | 4.356 | 1,059 |
| Food Products | 0.1381 | 0.0694 | 0.571 | 5.971 | 4,457 |
| Candy & Soda | 0.1722 | 0.0917 | 0.544 | 2.879 | 836 |
| Beer & Liquor | 0.1202 | 0.0595 | 0.585 | 5.911 | 1,048 |
| Tobacco Products | 0.1745 | 0.1425 | 0.482 | 2.038 | 412 |
| Recreation | 0.0816 | 0.0000 | 0.763 | 4.693 | 2,670 |
| Entertainment | 0.1078 | 0.0000 | 0.897 | 6.559 | 3,863 |
| Printing & Publishing | 0.1346 | 0.0745 | 0.683 | 12.526 | 2,899 |
| Consumer Goods | 0.1278 | 0.0418 | 0.728 | 10.937 | 5,624 |
| Apparel | 0.1054 | 0.0251 | 0.622 | 3.159 | 3,660 |
| Healthcare | 0.1677 | 0.0000 | 1.061 | 11.586 | 5,111 |
| Medical Equipment | 0.1299 | 0.0000 | 0.928 | 8.012 | 7,371 |
| Pharmaceutical Products | 0.1458 | -0.0204 | 1.001 | 5.382 | 11,530 |
| Chemicals | 0.1388 | 0.0668 | 0.593 | 3.422 | 4,830 |
| Rubber and Plastic Products | 0.1268 | 0.0306 | 0.625 | 2.831 | 2,364 |
| Textiles | 0.1126 | 0.0153 | 0.660 | 3.270 | 2,256 |
| Construction Materials | 0.1373 | 0.0524 | 0.593 | 3.336 | 7,202 |
| Construction | 0.1084 | 0.0029 | 0.666 | 2.514 | 3,355 |
| Steel Works | 0.1168 | 0.0297 | 0.579 | 3.129 | 3,915 |
| Fabricated Products | 0.1170 | 0.0333 | 0.572 | 2.021 | 1,169 |
| Machinery | 0.1412 | 0.0458 | 0.669 | 5.376 | 8,966 |
| Electrical Equipment | 0.1441 | -0.0000 | 1.101 | 10.144 | 6,648 |
| Automobiles and Trucks | 0.1202 | 0.0249 | 0.664 | 4.741 | 3,873 |
| Aircraft | 0.1599 | 0.0926 | 0.571 | 3.247 | 1,426 |
| Shipbuilding, Railroad Equipment | 0.1344 | 0.0690 | 0.602 | 4.471 | 492 |
| Defense | 0.2202 | 0.0854 | 1.134 | 11.263 | 473 |
| Precious Metals | 0.0897 | -0.0551 | 1.204 | 22.088 | 3,224 |
| Mines | 0.0594 | -0.0283 | 0.712 | 4.031 | 2,391 |
| Coal | 0.0758 | 0.0003 | 0.573 | 0.967 | 526 |
| Petroleum and Natural Gas | 0.0968 | -0.0000 | 0.795 | 5.108 | 13,825 |
| Utilities | 0.1371 | 0.1194 | 0.329 | 3.023 | 8,219 |
| Communication | 0.1586 | 0.0225 | 0.941 | 5.546 | 7,070 |
| Personal Services | 0.1133 | 0.0087 | 0.844 | 9.335 | 3,039 |
| Business Services | 0.1236 | 0.0040 | 0.767 | 4.066 | 16,689 |
| Computers | 0.1379 | -0.0204 | 1.015 | 9.349 | 7,595 |
| Computer Software | 0.1586 | -0.0006 | 1.148 | 8.633 | 13,625 |
| Electronic Equipment | 0.1700 | 0.0000 | 0.935 | 6.314 | 13,697 |
| Measuring and Control Equipment | 0.1716 | 0.0085 | 1.017 | 11.870 | 5,492 |
| Business Supplies | 0.1328 | 0.0704 | 0.572 | 5.657 | 2,733 |
| Shipping Containers | 0.1371 | 0.0729 | 0.507 | 2.856 | 1,427 |
| Transportation | 0.1133 | 0.0364 | 0.667 | 5.621 | 6,698 |
| Wholesale | 0.1176 | 0.0049 | 0.751 | 4.798 | 11,156 |
| Retail | 0.1333 | 0.0276 | 0.733 | 5.072 | 13,918 |
| Restaurant, Hotels, Motels | 0.1219 | 0.0175 | 1.412 | 53.031 | 5,924 |
| Banking | 0.1250 | 0.0798 | 0.481 | 3.785 | 22,401 |
| Insurance | 0.1374 | 0.0900 | 0.506 | 4.342 | 8,235 |
| Real Estate | 0.0909 | 0.0635 | 0.512 | 8.109 | 8,258 |
| Trading | 0.1624 | 0.0897 | 0.950 | 76.820 | 20,736 |
| Other | 0.0899 | 0.0000 | 0.707 | 4.231 | 1,138 |
| Not available | 0.0398 | 0.0000 | 0.680 | 5.237 | 1,946 |
| Total | 0.1321 | 0.0323 | 0.816 | 20.867 | 297,471 |

All common stocks and real estate investment trusts (REITs), occurring at some point in time, in CRSP from January 1970 until December 2018 are included. Stocks are assigned to an industry based on their Standard Industry Classification (SIC) code. Kenneth French's data library is used to structure which SIC codes belong to which industry. Companies are allowed to switch industries multiple times. If companies switch back to an industry they were previously in, this is treated as a separate run of returns to prevent faulty linking of returns. Mean and median returns depicted in the table are annual buy-and-hold returns including dividends and delisting returns.

Table 27: CRSP stocks buy-and-hold returns at decade horizon grouped by industry during 1970-2018

| | **Decade buy-and-hold returns** | | | | **Time in sample** | | |
|---|---|---|---|---|---|---|---|
| | Mean | Median | SD | Skewness | Mean (months) | SD | Obs |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 0.51 | -0.20 | 2.8 | 5.9 | 51 | 39 | 222 |
| Food Products | 1.14 | 0.34 | 2.9 | 6.0 | 65 | 42 | 773 |
| Candy & Soda | 1.56 | 0.45 | 3.2 | 2.4 | 67 | 42 | 140 |
| Beer & Liquor | 1.49 | 0.33 | 7.2 | 10.8 | 64 | 41 | 183 |
| Tobacco Products | 2.15 | 0.97 | 3.6 | 1.9 | 67 | 44 | 69 |
| Recreation | 0.71 | -0.36 | 7.0 | 14.6 | 58 | 40 | 505 |
| Entertainment | 0.60 | -0.23 | 3.3 | 9.9 | 49 | 37 | 846 |
| Printing & Publishing | 1.34 | 0.20 | 3.8 | 5.7 | 66 | 41 | 489 |
| Consumer Goods | 1.29 | 0.01 | 5.8 | 12.6 | 64 | 41 | 979 |
| Apparel | 0.84 | -0.01 | 2.9 | 5.2 | 67 | 40 | 612 |
| Healthcare | 1.07 | -0.09 | 5.3 | 12.0 | 49 | 36 | 1,102 |
| Medical Equipment | 0.75 | -0.13 | 3.1 | 5.7 | 56 | 40 | 1,428 |
| Pharmaceutical Products | 1.02 | -0.22 | 8.5 | 28.6 | 58 | 40 | 2,188 |
| Chemicals | 1.22 | 0.24 | 3.1 | 4.9 | 64 | 42 | 839 |
| Rubber and Plastic Products | 1.00 | 0.06 | 3.4 | 5.4 | 58 | 39 | 449 |
| Textiles | 0.82 | -0.06 | 2.8 | 3.6 | 63 | 40 | 398 |
| Construction Materials | 0.97 | 0.25 | 2.5 | 4.5 | 65 | 40 | 1,244 |
| Construction | 0.62 | -0.13 | 3.1 | 9.3 | 54 | 41 | 673 |
| Steel Works | 0.82 | 0.21 | 2.7 | 6.5 | 68 | 40 | 651 |
| Fabricated Products | 0.82 | 0.14 | 3.3 | 9.0 | 58 | 40 | 220 |
| Machinery | 0.92 | 0.18 | 2.5 | 4.0 | 65 | 42 | 1,557 |
| Electrical Equipment | 0.70 | -0.19 | 5.2 | 20.4 | 52 | 39 | 1,382 |
| Automobiles and Trucks | 0.87 | 0.04 | 2.8 | 5.0 | 67 | 42 | 647 |
| Aircraft | 1.32 | 0.46 | 2.7 | 2.9 | 69 | 41 | 233 |
| Shipbuilding, Railroad Equipment | 0.88 | 0.12 | 2.2 | 2.7 | 59 | 41 | 92 |
| Defense | 1.41 | 0.55 | 2.9 | 2.7 | 58 | 42 | 90 |
| Precious Metals | 0.19 | -0.48 | 2.4 | 6.7 | 56 | 40 | 621 |
| Mines | 0.24 | -0.26 | 1.9 | 4.4 | 54 | 39 | 477 |
| Coal | 0.54 | 0.00 | 2.1 | 3.6 | 53 | 41 | 108 |
| Petroleum and Natural Gas | 0.52 | -0.19 | 2.6 | 6.1 | 62 | 39 | 2,443 |
| Utilities | 1.64 | 0.97 | 2.0 | 1.7 | 85 | 41 | 1,120 |
| Communication | 1.15 | 0.00 | 6.7 | 16.3 | 50 | 38 | 1,508 |
| Personal Services | 0.81 | -0.15 | 4.5 | 10.9 | 50 | 37 | 648 |
| Business Services | 0.80 | -0.06 | 4.3 | 13.6 | 51 | 40 | 3,511 |
| Computers | 0.96 | -0.25 | 12.0 | 25.3 | 54 | 39 | 1,529 |
| Computer Software | 0.72 | -0.32 | 7.1 | 34.9 | 48 | 36 | 2,994 |
| Electronic Equipment | 1.02 | -0.06 | 3.9 | 6.9 | 63 | 41 | 2,409 |
| Measuring and Control Equipment | 0.91 | 0.06 | 3.4 | 6.9 | 61 | 41 | 991 |
| Business Supplies | 1.25 | 0.37 | 3.3 | 5.6 | 65 | 41 | 472 |
| Shipping Containers | 1.15 | 0.26 | 2.5 | 3.1 | 62 | 40 | 257 |
| Transportation | 0.95 | 0.08 | 4.6 | 18.0 | 62 | 41 | 1,210 |
| Wholesale | 0.70 | -0.06 | 3.4 | 14.4 | 52 | 39 | 2,303 |
| Retail | 1.09 | 0.00 | 4.2 | 7.5 | 61 | 40 | 2,522 |
| Restaurant, Hotels, Motels | 0.91 | -0.04 | 4.5 | 12.4 | 60 | 39 | 1,087 |
| Banking | 0.94 | 0.32 | 2.7 | 11.6 | 54 | 39 | 4,439 |
| Insurance | 1.22 | 0.35 | 3.1 | 6.5 | 64 | 41 | 1,446 |
| Real Estate | 0.72 | 0.12 | 2.4 | 7.2 | 57 | 39 | 1,580 |
| Trading | 1.22 | 0.19 | 4.0 | 13.6 | 47 | 38 | 4,662 |
| Other | 0.81 | -0.14 | 3.8 | 6.4 | 48 | 39 | 251 |
| Not available | 0.15 | 0.00 | 1.9 | 9.4 | 27 | 30 | 657 |
| Total | 0.92 | 0.04 | 4.7 | 30.9 | 57 | 40 | 57,256 |

All common stocks and real estate investment trusts (REITs), occurring at some point in time, in CRSP from January 1970 until December 2018 are included. Stocks are assigned to an industry based on their Standard Industry Classification (SIC) code. Kenneth French's data library is used to structure which SIC codes belong to which industry. Companies are allowed to switch industries multiple times. If companies switch back to an industry they were previously in, this is treated as a separate run of returns to prevent faulty linking of returns. Mean and median returns depicted in the table are decade buy-and-hold returns including dividends and delisting returns.

Table 28: CRSP stocks buy-and-hold returns at lifetime horizon grouped by industry during 1970-2018

| | **Lifetime buy-and-hold returns** | | | | **Time in sample** | | |
|---|---|---|---|---|---|---|---|
| | Mean | Median | SD | Skewness | Mean (months) | SD | Obs |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 1.78 | -0.35 | 8.7 | 6.7 | 85 | 87 | 133 |
| Food Products | 14.33 | 0.52 | 68.6 | 8.1 | 122 | 130 | 409 |
| Candy & Soda | 26.75 | 0.52 | 97.5 | 4.1 | 128 | 132 | 74 |
| Beer & Liquor | 7.59 | 0.19 | 26.6 | 4.4 | 124 | 121 | 95 |
| Tobacco Products | 49.59 | 0.79 | 194.3 | 5.3 | 129 | 125 | 36 |
| Recreation | 3.77 | -0.56 | 47.2 | 16.6 | 95 | 99 | 307 |
| Entertainment | 2.64 | -0.43 | 18.4 | 9.6 | 77 | 75 | 531 |
| Printing & Publishing | 12.62 | 0.13 | 88.0 | 13.6 | 123 | 129 | 263 |
| Consumer Goods | 6.56 | -0.07 | 36.6 | 9.6 | 120 | 118 | 521 |
| Apparel | 11.41 | -0.32 | 98.8 | 12.2 | 127 | 119 | 324 |
| Healthcare | 3.10 | -0.13 | 25.1 | 15.7 | 78 | 75 | 693 |
| Medical Equipment | 6.89 | -0.28 | 89.7 | 25.5 | 97 | 95 | 828 |
| Pharmaceutical Products | 5.29 | -0.43 | 47.8 | 15.9 | 95 | 94 | 1,339 |
| Chemicals | 9.40 | 0.36 | 46.5 | 9.3 | 117 | 119 | 460 |
| Rubber and Plastic Products | 4.20 | 0.04 | 18.4 | 6.0 | 100 | 92 | 259 |
| Textiles | 1.96 | 0.02 | 5.3 | 3.0 | 118 | 102 | 214 |
| Construction Materials | 5.49 | 0.34 | 27.7 | 12.3 | 123 | 113 | 659 |
| Construction | 1.70 | -0.17 | 14.7 | 16.3 | 91 | 97 | 400 |
| Steel Works | 15.25 | 0.09 | 204.3 | 18.2 | 129 | 113 | 342 |
| Fabricated Products | 2.96 | 0.08 | 11.2 | 5.4 | 102 | 87 | 125 |
| Machinery | 8.70 | 0.13 | 52.6 | 13.2 | 123 | 124 | 814 |
| Electrical Equipment | 2.28 | -0.44 | 15.6 | 12.0 | 92 | 86 | 785 |
| Automobiles and Trucks | 5.27 | -0.02 | 29.1 | 10.1 | 127 | 123 | 343 |
| Aircraft | 19.66 | 0.60 | 115.7 | 9.4 | 129 | 117 | 125 |
| Shipbuilding, Railroad Equipment | 6.90 | 0.12 | 40.3 | 7.1 | 101 | 106 | 54 |
| Defense | 17.80 | 0.34 | 95.8 | 6.9 | 99 | 108 | 53 |
| Precious Metals | 0.60 | -0.71 | 6.7 | 9.9 | 98 | 89 | 356 |
| Mines | 0.75 | -0.47 | 6.4 | 12.0 | 83 | 90 | 308 |
| Coal | 1.46 | 0.05 | 5.1 | 3.5 | 94 | 103 | 61 |
| Petroleum and Natural Gas | 6.56 | -0.55 | 139.6 | 36.6 | 104 | 109 | 1,454 |
| Utilities | 44.81 | 3.68 | 140.2 | 8.3 | 224 | 190 | 426 |
| Communication | 3.93 | -0.18 | 24.3 | 10.7 | 81 | 79 | 926 |
| Personal Services | 1.50 | -0.22 | 7.6 | 7.3 | 81 | 80 | 402 |
| Business Services | 3.18 | -0.18 | 27.1 | 15.1 | 81 | 86 | 2,201 |
| Computers | 2.57 | -0.47 | 31.0 | 22.1 | 96 | 88 | 865 |
| Computer Software | 4.22 | -0.50 | 56.8 | 21.2 | 73 | 71 | 1,955 |
| Electronic Equipment | 5.65 | -0.25 | 55.4 | 25.6 | 114 | 109 | 1,334 |
| Measuring and Control Equipment | 3.13 | 0.00 | 14.9 | 8.5 | 109 | 104 | 555 |
| Business Supplies | 8.29 | 0.52 | 32.9 | 6.4 | 128 | 123 | 239 |
| Shipping Containers | 6.63 | 0.59 | 41.2 | 10.9 | 109 | 107 | 147 |
| Transportation | 20.14 | -0.07 | 353.3 | 24.3 | 111 | 108 | 671 |
| Wholesale | 3.99 | -0.22 | 35.5 | 17.0 | 88 | 96 | 1,369 |
| Retail | 15.73 | -0.17 | 218.3 | 27.8 | 109 | 106 | 1,407 |
| Restaurant, Hotels, Motels | 5.06 | -0.25 | 47.2 | 18.2 | 105 | 102 | 620 |
| Banking | 2.32 | 0.64 | 8.7 | 11.0 | 89 | 79 | 2,727 |
| Insurance | 9.57 | 0.59 | 73.1 | 20.5 | 122 | 115 | 758 |
| Real Estate | 2.18 | 0.13 | 8.5 | 7.3 | 96 | 94 | 939 |
| Trading | 4.91 | 0.53 | 51.0 | 38.4 | 81 | 76 | 2,743 |
| Other | 1.24 | -0.25 | 5.4 | 5.2 | 77 | 76 | 156 |
| Not available | 0.75 | -0.01 | 11.2 | 18.8 | 32 | 47 | 552 |
| Total | 6.17 | -0.01 | 88.1 | 61.6 | 97 | 100 | 33,357 |

All common stocks and real estate investment trusts (REITs), occurring at some point in time, in CRSP from January 1970 until December 2018 are included. Stocks are assigned to an industry based on their Standard Industry Classification (SIC) code. Kenneth French's data library is used to structure which SIC codes belong to which industry. Companies are allowed to switch industries multiple times. If companies switch back to an industry they were previously in, this is treated as a separate run of returns to prevent faulty linking of returns. Mean and median returns depicted in the table are life buy-and-hold returns including dividends and delisting returns.

# Appendix B: variability in corporate performance

Table 29: Mean standard deviation of firm financial ratios per industry during 1970-2018

| | Mean standard deviation across firms per industry | | | | | | |
|---|---|---|---|---|---|---|---|
| | **Capitalization** | **Efficiency** | **Liquidity** | **Profitability** | | **Solvency** | |
| | Capital ratio | Asset turnover | Quick ratio | Return on assets | Operating margin | Cash flow/debt | Interest coverage |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 0.10 | 0.19 | 0.97 | 6.8% | 59.1% | 0.30 | 13.48 |
| Food Products | 0.11 | 0.29 | 0.57 | 6.5% | 12.8% | 0.20 | 17.26 |
| Candy & Soda | 0.09 | 0.22 | 0.42 | 6.0% | 3.3% | 0.15 | 24.45 |
| Beer & Liquor | 0.09 | 0.23 | 0.54 | 6.7% | 25.9% | 0.22 | 20.89 |
| Tobacco Products | 0.12 | 0.20 | 0.77 | 8.0% | 65.9% | 0.27 | 19.76 |
| Recreation | 0.11 | 0.27 | 0.76 | 9.4% | 25.0% | 0.32 | 33.35 |
| Entertainment | 0.13 | 0.22 | 0.99 | 10.1% | 55.0% | 0.33 | 28.26 |
| Printing & Publishing | 0.12 | 0.20 | 0.73 | 7.3% | 20.2% | 0.21 | 30.76 |
| Consumer Goods | 0.10 | 0.25 | 0.56 | 7.6% | 17.5% | 0.22 | 23.57 |
| Apparel | 0.10 | 0.28 | 0.67 | 8.8% | 12.6% | 0.27 | 23.17 |
| Healthcare | 0.11 | 0.28 | 1.12 | 8.3% | 33.7% | 0.28 | 33.68 |
| Medical Equipment | 0.11 | 0.21 | 1.69 | 11.6% | 96.1% | 0.53 | 65.02 |
| Pharmaceutical Products | 0.11 | 0.18 | 2.53 | 14.6% | 231.6% | 0.72 | 78.96 |
| Chemicals | 0.11 | 0.22 | 0.98 | 7.7% | 41.5% | 0.31 | 29.97 |
| Rubber and Plastic Products | 0.11 | 0.22 | 0.50 | 6.6% | 18.4% | 0.17 | 16.76 |
| Textiles | 0.08 | 0.23 | 0.41 | 6.8% | 4.5% | 0.17 | 12.22 |
| Construction Materials | 0.10 | 0.25 | 0.55 | 7.4% | 15.0% | 0.19 | 18.07 |
| Construction | 0.10 | 0.28 | 0.52 | 7.2% | 16.2% | 0.19 | 18.32 |
| Steel Works | 0.11 | 0.26 | 0.45 | 6.7% | 6.3% | 0.16 | 17.89 |
| Fabricated Products | 0.08 | 0.23 | 0.32 | 6.7% | 4.9% | 0.16 | 16.53 |
| Machinery | 0.10 | 0.21 | 0.64 | 7.7% | 22.6% | 0.23 | 29.36 |
| Electrical Equipment | 0.10 | 0.25 | 1.05 | 10.6% | 42.5% | 0.36 | 50.33 |
| Automobiles and Trucks | 0.11 | 0.26 | 0.45 | 7.2% | 20.3% | 0.19 | 19.47 |
| Aircraft | 0.12 | 0.24 | 0.44 | 6.8% | 9.7% | 0.18 | 16.48 |
| Shipbuilding, Railroad Equipment | 0.12 | 0.31 | 0.41 | 7.0% | 6.0% | 0.16 | 9.95 |
| Defense | 0.11 | 0.18 | 0.48 | 5.3% | 16.0% | 0.22 | 23.82 |
| Precious Metals | 0.09 | 0.12 | 2.00 | 8.8% | 175.9% | 0.61 | 38.35 |
| Mines | 0.10 | 0.16 | 1.58 | 8.2% | 88.5% | 0.41 | 27.25 |
| Coal | 0.15 | 0.22 | 0.82 | 7.8% | 72.1% | 0.18 | 13.01 |
| Petroleum and Natural Gas | 0.13 | 0.16 | 1.08 | 9.1% | 77.4% | 0.32 | 25.09 |
| Utilities | 0.07 | 0.16 | 0.43 | 3.2% | 18.2% | 0.07 | 6.23 |
| Communication | 0.14 | 0.18 | 1.03 | 7.3% | 50.5% | 0.22 | 19.45 |
| Personal Services | 0.10 | 0.23 | 0.70 | 7.8% | 23.1% | 0.23 | 31.39 |
| Business Services | 0.10 | 0.25 | 1.13 | 8.8% | 53.1% | 0.32 | 40.66 |
| Computers | 0.11 | 0.28 | 1.00 | 12.4% | 46.8% | 0.40 | 51.57 |
| Computer Software | 0.09 | 0.26 | 1.22 | 12.2% | 55.2% | 0.40 | 74.59 |
| Electronic Equipment | 0.10 | 0.25 | 1.03 | 10.4% | 33.8% | 0.35 | 49.18 |
| Measuring and Control Equipment | 0.09 | 0.22 | 1.07 | 9.4% | 46.4% | 0.38 | 49.69 |
| Business Supplies | 0.10 | 0.19 | 0.44 | 5.6% | 11.1% | 0.16 | 17.22 |
| Shipping Containers | 0.09 | 0.20 | 0.41 | 6.5% | 12.4% | 0.17 | 14.38 |
| Transportation | 0.11 | 0.27 | 0.50 | 6.4% | 11.1% | 0.14 | 18.06 |
| Wholesale | 0.10 | 0.31 | 0.59 | 7.7% | 21.6% | 0.25 | 26.82 |
| Retail | 0.11 | 0.31 | 0.46 | 7.0% | 12.6% | 0.18 | 27.22 |
| Restaurant, Hotels, Motels | 0.12 | 0.26 | 0.67 | 6.2% | 9.7% | 0.18 | 18.28 |
| Banking | 0.10 | 0.03 | 0.94 | 1.6% | 11.9% | 0.08 | 16.27 |
| Insurance | 0.08 | 0.13 | 0.74 | 3.7% | 16.2% | 0.12 | 37.94 |
| Real Estate | 0.10 | 0.16 | 0.77 | 5.0% | 31.1% | 0.19 | 7.87 |
| Trading | 0.09 | 0.10 | 1.09 | 3.4% | 17.6% | 0.14 | 25.24 |
| Other | 0.11 | 0.20 | 0.74 | 7.3% | 40.8% | 0.18 | 15.57 |
| Not available | 0.07 | 0.21 | 1.18 | 7.3% | 40.0% | 0.27 | 29.03 |
| Total | 0.10 | 0.22 | 0.82 | 7.6% | 37.2% | 0.25 | 27.44 |

The table displays the mean standard deviation of financial ratios for every industry during 1970-2018.

# Appendix C: benchmarked stock performance

Table 30: Percentage of stocks that outperform their benchmark over a monthly horizon during 1970-2018

| | **Non-stock benchmarks** | | | **Market benchmarks** | | **Industry benchmarks** | |
|---|---|---|---|---|---|---|---|
| | % > 0 | % > T-bill | % > CPI | % > EW Mkt return | % > VW Mkt return | % > EW Ind return | % > VW Ind return |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 42 | 41 | 42 | 43 | 44 | 46 | 44 |
| Food Products | 49 | 48 | 48 | 47 | 47 | 46 | 47 |
| Candy & Soda | 52 | 51 | 51 | 48 | 49 | 47 | 49 |
| Beer & Liquor | 48 | 47 | 47 | 46 | 46 | 47 | 46 |
| Tobacco Products | 53 | 51 | 52 | 49 | 50 | 48 | 47 |
| Recreation | 42 | 42 | 43 | 44 | 44 | 46 | 46 |
| Entertainment | 43 | 43 | 44 | 44 | 44 | 46 | 45 |
| Printing & Publishing | 49 | 48 | 48 | 47 | 47 | 47 | 47 |
| Consumer Goods | 46 | 45 | 46 | 46 | 46 | 46 | 47 |
| Apparel | 45 | 45 | 45 | 45 | 46 | 47 | 46 |
| Healthcare | 46 | 46 | 46 | 46 | 46 | 46 | 47 |
| Medical Equipment | 45 | 45 | 46 | 45 | 45 | 46 | 45 |
| Pharmaceutical Products | 45 | 45 | 46 | 45 | 45 | 45 | 45 |
| Chemicals | 49 | 48 | 49 | 47 | 48 | 47 | 47 |
| Rubber and Plastic Products | 46 | 45 | 46 | 46 | 46 | 46 | 47 |
| Textiles | 45 | 44 | 45 | 45 | 46 | 47 | 46 |
| Construction Materials | 46 | 45 | 46 | 46 | 46 | 46 | 47 |
| Construction | 45 | 45 | 45 | 45 | 45 | 46 | 47 |
| Steel Works | 47 | 46 | 47 | 46 | 46 | 47 | 48 |
| Fabricated Products | 45 | 44 | 45 | 45 | 45 | 46 | 47 |
| Machinery | 48 | 47 | 47 | 46 | 47 | 46 | 48 |
| Electrical Equipment | 44 | 44 | 44 | 44 | 44 | 45 | 45 |
| Automobiles and Trucks | 47 | 46 | 47 | 46 | 46 | 47 | 47 |
| Aircraft | 48 | 47 | 48 | 47 | 48 | 46 | 47 |
| Shipbuilding, Railroad Equipment | 49 | 48 | 48 | 48 | 48 | 49 | 47 |
| Defense | 51 | 50 | 50 | 48 | 48 | 47 | 48 |
| Precious Metals | 40 | 40 | 40 | 42 | 42 | 46 | 45 |
| Mines | 42 | 41 | 42 | 43 | 43 | 46 | 45 |
| Coal | 47 | 46 | 46 | 45 | 46 | 48 | 48 |
| Petroleum and Natural Gas | 43 | 43 | 43 | 44 | 44 | 46 | 45 |
| Utilities | 56 | 54 | 54 | 49 | 50 | 48 | 50 |
| Communication | 48 | 47 | 47 | 46 | 46 | 46 | 46 |
| Personal Services | 46 | 45 | 46 | 45 | 46 | 47 | 47 |
| Business Services | 46 | 46 | 46 | 46 | 46 | 46 | 46 |
| Computers | 44 | 44 | 44 | 44 | 44 | 46 | 46 |
| Computer Software | 46 | 46 | 46 | 45 | 45 | 45 | 45 |
| Electronic Equipment | 45 | 45 | 46 | 46 | 46 | 46 | 46 |
| Measuring and Control Equipment | 45 | 45 | 45 | 45 | 45 | 45 | 46 |
| Business Supplies | 49 | 48 | 49 | 47 | 47 | 47 | 48 |
| Shipping Containers | 49 | 48 | 48 | 47 | 47 | 47 | 47 |
| Transportation | 48 | 47 | 48 | 46 | 46 | 47 | 47 |
| Wholesale | 45 | 45 | 45 | 45 | 45 | 46 | 46 |
| Retail | 47 | 46 | 47 | 46 | 46 | 47 | 47 |
| Restaurant, Hotels, Motels | 45 | 45 | 45 | 45 | 45 | 46 | 46 |
| Banking | 51 | 50 | 51 | 48 | 48 | 47 | 49 |
| Insurance | 51 | 50 | 50 | 48 | 48 | 47 | 48 |
| Real Estate | 49 | 49 | 49 | 47 | 47 | 48 | 47 |
| Trading | 50 | 49 | 49 | 47 | 47 | 46 | 46 |
| Other | 46 | 46 | 46 | 45 | 45 | 46 | 46 |
| Not available | 45 | 44 | 44 | 43 | 43 | 47 | 45 |
| Total | 46 | 46 | 46 | 46 | 46 | 46 | 47 |

All mdate buy-and-hold common stock and REIT stock returns from January 1970 until December 2018 are benchmarked against zero, 30-day Treasury Bills, the Consumer Price Index (CPI), equal- and value-weighted market returns and equal- and value-weighted industry returns. Equal-weighted (EW) means that all stocks are given the same weight when calculating market and industry returns. Value-weighted (VW) means that stocks are given a weight based on their market capitalization when calculating market and industry returns. Stocks are allowed to switch industries multiple times. If stocks switch back to the industry they were previously in, this is treated as a seperate run of returns to prevent faulty linking of returns. Buy-and-hold stock returns are exactly matched to the benchmark on a monthly frequency.

Table 31: Percentage of stocks that outperform their benchmark over an annual horizon during 1926-2018

| | **Non-stock benchmarks** | | | **Market benchmarks** | | **Industry benchmarks** | |
|---|---|---|---|---|---|---|---|
| | % > 0 | % > T-bill | % > CPI | % > EW Mkt return | % > VW Mkt return | % > EW Ind return | % > VW Ind return |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 49 | 43 | 44 | 38 | 38 | 41 | 40 |
| Food Products | 59 | 55 | 55 | 44 | 46 | 43 | 46 |
| Candy & Soda | 67 | 60 | 60 | 46 | 50 | 42 | 51 |
| Beer & Liquor | 58 | 54 | 54 | 43 | 44 | 45 | 45 |
| Tobacco Products | 64 | 60 | 60 | 45 | 50 | 43 | 49 |
| Recreation | 47 | 44 | 45 | 37 | 39 | 42 | 42 |
| Entertainment | 49 | 46 | 47 | 39 | 41 | 41 | 40 |
| Printing & Publishing | 59 | 54 | 55 | 45 | 46 | 42 | 46 |
| Consumer Goods | 55 | 51 | 51 | 42 | 44 | 43 | 45 |
| Apparel | 54 | 50 | 51 | 40 | 44 | 43 | 45 |
| Healthcare | 50 | 47 | 48 | 43 | 43 | 41 | 44 |
| Medical Equipment | 48 | 46 | 46 | 41 | 41 | 41 | 40 |
| Pharmaceutical Products | 47 | 45 | 45 | 40 | 41 | 39 | 41 |
| Chemicals | 59 | 55 | 55 | 44 | 46 | 43 | 47 |
| Rubber and Plastic Products | 54 | 49 | 50 | 42 | 44 | 42 | 45 |
| Textiles | 54 | 50 | 50 | 41 | 43 | 44 | 44 |
| Construction Materials | 57 | 52 | 53 | 42 | 45 | 42 | 45 |
| Construction | 51 | 48 | 48 | 41 | 42 | 43 | 44 |
| Steel Works | 57 | 53 | 53 | 42 | 45 | 42 | 47 |
| Fabricated Products | 53 | 48 | 49 | 41 | 43 | 42 | 45 |
| Machinery | 57 | 53 | 54 | 43 | 45 | 42 | 47 |
| Electrical Equipment | 49 | 46 | 47 | 40 | 41 | 41 | 41 |
| Automobiles and Trucks | 54 | 51 | 52 | 42 | 44 | 43 | 44 |
| Aircraft | 59 | 56 | 56 | 46 | 49 | 42 | 46 |
| Shipbuilding, Railroad Equipment | 58 | 55 | 56 | 42 | 45 | 43 | 47 |
| Defense | 59 | 54 | 55 | 47 | 49 | 46 | 45 |
| Precious Metals | 42 | 39 | 40 | 36 | 35 | 38 | 40 |
| Mines | 48 | 44 | 45 | 37 | 39 | 40 | 43 |
| Coal | 53 | 49 | 50 | 41 | 44 | 45 | 47 |
| Petroleum and Natural Gas | 49 | 46 | 46 | 39 | 40 | 41 | 40 |
| Utilities | 71 | 63 | 64 | 47 | 49 | 46 | 51 |
| Communication | 53 | 51 | 51 | 44 | 44 | 43 | 45 |
| Personal Services | 51 | 48 | 48 | 41 | 43 | 43 | 46 |
| Business Services | 51 | 48 | 48 | 42 | 43 | 41 | 43 |
| Computers | 47 | 44 | 44 | 38 | 39 | 40 | 40 |
| Computer Software | 48 | 46 | 46 | 41 | 42 | 40 | 40 |
| Electronic Equipment | 50 | 47 | 48 | 41 | 42 | 40 | 43 |
| Measuring and Control Equipment | 51 | 48 | 49 | 41 | 43 | 39 | 42 |
| Business Supplies | 60 | 55 | 56 | 44 | 47 | 44 | 49 |
| Shipping Containers | 61 | 56 | 56 | 45 | 49 | 44 | 48 |
| Transportation | 55 | 51 | 52 | 42 | 43 | 44 | 45 |
| Wholesale | 51 | 48 | 48 | 41 | 42 | 42 | 42 |
| Retail | 54 | 50 | 51 | 42 | 44 | 43 | 45 |
| Restaurant, Hotels, Motels | 52 | 48 | 49 | 41 | 43 | 43 | 43 |
| Banking | 62 | 57 | 58 | 48 | 49 | 45 | 50 |
| Insurance | 62 | 57 | 58 | 47 | 48 | 44 | 47 |
| Real Estate | 59 | 54 | 55 | 44 | 45 | 46 | 46 |
| Trading | 62 | 56 | 57 | 47 | 48 | 44 | 46 |
| Other | 51 | 48 | 49 | 42 | 41 | 43 | 45 |
| Not available | 50 | 46 | 46 | 39 | 38 | 44 | 43 |
| Total | 54 | 50 | 50 | 42 | 44 | 43 | 45 |

All annual buy-and-hold common stock and REIT stock returns from January 1926 until December 2018 are benchmarked against zero, 30-day Treasury Bills, the Consumer Price Index (CPI), equal- and value-weighted market returns and equal- and value-weighted industry returns. Equal-weighted (EW) means that all stocks are given the same weight when calculating market and industry returns. Value-weighted (VW) means that stocks are given a weight based on their market capitalization when calculating market and industry returns. Stocks are allowed to switch industries multiple times. If stocks switch back to the industry they were previously in, this is treated as a seperate run of returns to prevent faulty linking of returns. Buy-and-hold stock returns are exactly matched to the benchmark on a monthly frequency.

Table 32: Percentage of stocks that outperform their benchmark over an annual horizon during 1970-2018

| | Non-stock benchmarks | | | Market benchmarks | | Industry benchmarks | |
|---|---|---|---|---|---|---|---|
| | % > 0 | % > T-bill | % > CPI | % > EW Mkt return | % > VW Mkt return | % > EW Ind return | % > VW Ind return |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 47 | 42 | 43 | 37 | 37 | 41 | 39 |
| Food Products | 58 | 53 | 53 | 46 | 47 | 42 | 44 |
| Candy & Soda | 66 | 59 | 59 | 47 | 51 | 42 | 53 |
| Beer & Liquor | 57 | 52 | 52 | 43 | 45 | 45 | 45 |
| Tobacco Products | 67 | 61 | 60 | 50 | 56 | 46 | 49 |
| Recreation | 47 | 43 | 44 | 37 | 39 | 42 | 41 |
| Entertainment | 47 | 44 | 45 | 39 | 40 | 41 | 39 |
| Printing & Publishing | 59 | 54 | 54 | 44 | 46 | 42 | 46 |
| Consumer Goods | 54 | 49 | 50 | 42 | 43 | 43 | 45 |
| Apparel | 53 | 48 | 49 | 40 | 44 | 43 | 43 |
| Healthcare | 50 | 47 | 48 | 43 | 43 | 41 | 44 |
| Medical Equipment | 47 | 45 | 45 | 41 | 41 | 41 | 39 |
| Pharmaceutical Products | 46 | 44 | 44 | 39 | 40 | 38 | 40 |
| Chemicals | 58 | 53 | 54 | 44 | 45 | 43 | 45 |
| Rubber and Plastic Products | 53 | 49 | 50 | 42 | 44 | 42 | 45 |
| Textiles | 52 | 47 | 48 | 41 | 42 | 44 | 43 |
| Construction Materials | 56 | 50 | 51 | 43 | 44 | 41 | 44 |
| Construction | 51 | 48 | 48 | 41 | 42 | 43 | 44 |
| Steel Works | 53 | 48 | 49 | 41 | 42 | 42 | 46 |
| Fabricated Products | 53 | 47 | 49 | 41 | 43 | 41 | 45 |
| Machinery | 55 | 51 | 51 | 43 | 44 | 42 | 46 |
| Electrical Equipment | 47 | 44 | 44 | 39 | 39 | 40 | 40 |
| Automobiles and Trucks | 52 | 48 | 49 | 41 | 43 | 44 | 45 |
| Aircraft | 59 | 55 | 56 | 48 | 50 | 43 | 45 |
| Shipbuilding, Railroad Equipment | 58 | 53 | 54 | 45 | 46 | 43 | 45 |
| Defense | 60 | 55 | 56 | 47 | 49 | 46 | 45 |
| Precious Metals | 41 | 38 | 39 | 35 | 34 | 38 | 40 |
| Mines | 43 | 39 | 40 | 34 | 35 | 40 | 40 |
| Coal | 51 | 45 | 46 | 41 | 41 | 46 | 47 |
| Petroleum and Natural Gas | 47 | 43 | 44 | 37 | 39 | 41 | 39 |
| Utilities | 71 | 63 | 64 | 47 | 51 | 47 | 52 |
| Communication | 53 | 50 | 50 | 43 | 44 | 43 | 44 |
| Personal Services | 51 | 47 | 48 | 41 | 42 | 43 | 45 |
| Business Services | 51 | 47 | 48 | 41 | 43 | 41 | 42 |
| Computers | 46 | 43 | 43 | 38 | 38 | 40 | 40 |
| Computer Software | 48 | 46 | 46 | 41 | 42 | 40 | 40 |
| Electronic Equipment | 50 | 47 | 47 | 41 | 42 | 40 | 43 |
| Measuring and Control Equipment | 51 | 47 | 48 | 41 | 42 | 39 | 42 |
| Business Supplies | 60 | 54 | 55 | 44 | 47 | 44 | 48 |
| Shipping Containers | 61 | 53 | 54 | 44 | 46 | 42 | 48 |
| Transportation | 54 | 50 | 50 | 43 | 44 | 45 | 45 |
| Wholesale | 51 | 47 | 48 | 41 | 42 | 42 | 42 |
| Retail | 53 | 49 | 49 | 42 | 44 | 43 | 44 |
| Restaurant, Hotels, Motels | 52 | 48 | 49 | 41 | 43 | 43 | 43 |
| Banking | 62 | 57 | 58 | 48 | 49 | 46 | 50 |
| Insurance | 62 | 57 | 58 | 47 | 48 | 44 | 48 |
| Real Estate | 59 | 54 | 55 | 45 | 45 | 46 | 46 |
| Trading | 62 | 56 | 57 | 48 | 48 | 44 | 46 |
| Other | 50 | 47 | 48 | 42 | 41 | 43 | 44 |
| Not available | 48 | 45 | 45 | 38 | 38 | 44 | 43 |
| Total | 53 | 48 | 49 | 42 | 43 | 43 | 44 |

All annual buy-and-hold common stock and REIT stock returns from January 1970 until December 2018 are benchmarked against zero, 30-day Treasury Bills, the Consumer Price Index (CPI), equal- and value-weighted market returns and equal- and value-weighted industry returns. Equal-weighted (EW) means that all stocks are given the same weight when calculating market and industry returns. Value-weighted (VW) means that stocks are given a weight based on their market capitalization when calculating market and industry returns. Stocks are allowed to switch industries multiple times. If stocks switch back to the industry they were previously in, this is treated as a seperate run of returns to prevent faulty linking of returns. Buy-and-hold stock returns are exactly matched to the benchmark on a monthly frequency.

Table 33: Percentage of stocks that outperform their benchmark over a decade horizon during 1926-2018

| | **Non-stock benchmarks** | | | **Market benchmarks** | | **Industry benchmarks** | |
|---|---|---|---|---|---|---|---|
| | % > 0 | % > T-bill | % > CPI | % > EW Mkt return | % > VW Mkt return | % > EW Ind return | % > VW Ind return |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 42 | 34 | 34 | 25 | 31 | 32 | 32 |
| Food Products | 68 | 60 | 60 | 37 | 45 | 34 | 42 |
| Candy & Soda | 73 | 65 | 62 | 42 | 53 | 35 | 48 |
| Beer & Liquor | 63 | 56 | 57 | 35 | 41 | 41 | 35 |
| Tobacco Products | 74 | 72 | 63 | 43 | 48 | 33 | 46 |
| Recreation | 41 | 34 | 36 | 22 | 29 | 30 | 33 |
| Entertainment | 44 | 39 | 40 | 30 | 32 | 31 | 30 |
| Printing & Publishing | 60 | 52 | 54 | 38 | 42 | 33 | 40 |
| Consumer Goods | 54 | 46 | 47 | 31 | 37 | 34 | 38 |
| Apparel | 53 | 45 | 47 | 30 | 37 | 38 | 39 |
| Healthcare | 47 | 41 | 42 | 33 | 34 | 32 | 38 |
| Medical Equipment | 47 | 42 | 42 | 32 | 33 | 31 | 30 |
| Pharmaceutical Products | 43 | 39 | 39 | 32 | 32 | 28 | 30 |
| Chemicals | 64 | 55 | 56 | 37 | 42 | 36 | 42 |
| Rubber and Plastic Products | 55 | 47 | 48 | 35 | 38 | 35 | 42 |
| Textiles | 53 | 46 | 47 | 28 | 35 | 36 | 38 |
| Construction Materials | 64 | 54 | 55 | 34 | 42 | 33 | 43 |
| Construction | 48 | 41 | 42 | 31 | 33 | 36 | 38 |
| Steel Works | 61 | 52 | 53 | 29 | 38 | 33 | 44 |
| Fabricated Products | 57 | 48 | 50 | 32 | 41 | 34 | 43 |
| Machinery | 61 | 53 | 54 | 35 | 42 | 34 | 45 |
| Electrical Equipment | 45 | 38 | 39 | 29 | 31 | 29 | 30 |
| Automobiles and Trucks | 58 | 51 | 52 | 30 | 37 | 34 | 39 |
| Aircraft | 67 | 61 | 59 | 39 | 47 | 32 | 41 |
| Shipbuilding, Railroad Equipment | 63 | 59 | 55 | 33 | 39 | 36 | 43 |
| Defense | 65 | 57 | 61 | 44 | 44 | 36 | 36 |
| Precious Metals | 31 | 26 | 27 | 22 | 22 | 28 | 29 |
| Mines | 45 | 37 | 38 | 27 | 29 | 31 | 35 |
| Coal | 56 | 50 | 50 | 34 | 36 | 38 | 40 |
| Petroleum and Natural Gas | 47 | 40 | 42 | 28 | 33 | 29 | 31 |
| Utilities | 86 | 73 | 72 | 42 | 50 | 37 | 56 |
| Communication | 52 | 47 | 48 | 38 | 39 | 33 | 38 |
| Personal Services | 45 | 41 | 42 | 32 | 34 | 35 | 39 |
| Business Services | 48 | 43 | 44 | 35 | 35 | 33 | 36 |
| Computers | 42 | 35 | 36 | 27 | 29 | 29 | 29 |
| Computer Software | 40 | 37 | 38 | 30 | 30 | 28 | 27 |
| Electronic Equipment | 49 | 41 | 43 | 31 | 33 | 28 | 36 |
| Measuring and Control Equipment | 54 | 44 | 47 | 32 | 36 | 27 | 33 |
| Business Supplies | 65 | 58 | 59 | 37 | 47 | 37 | 47 |
| Shipping Containers | 68 | 60 | 60 | 41 | 45 | 36 | 47 |
| Transportation | 55 | 49 | 50 | 31 | 36 | 36 | 38 |
| Wholesale | 48 | 43 | 44 | 33 | 35 | 34 | 34 |
| Retail | 54 | 47 | 48 | 33 | 38 | 36 | 38 |
| Restaurant, Hotels, Motels | 49 | 43 | 44 | 31 | 33 | 34 | 35 |
| Banking | 66 | 59 | 60 | 46 | 48 | 44 | 50 |
| Insurance | 66 | 57 | 58 | 40 | 46 | 38 | 46 |
| Real Estate | 56 | 50 | 53 | 37 | 35 | 40 | 41 |
| Trading | 64 | 55 | 58 | 45 | 44 | 40 | 40 |
| Other | 45 | 40 | 40 | 35 | 32 | 36 | 38 |
| Not available | 47 | 43 | 41 | 32 | 32 | 39 | 38 |
| Total | 54 | 47 | 48 | 33 | 37 | 34 | 38 |

All decade buy-and-hold common stock and REIT stock returns from January 1926 until December 2018 are benchmarked against zero, 30-day Treasury Bills, the Consumer Price Index (CPI), equal- and value-weighted market returns and equal- and value-weighted industry returns. Equal-weighted (EW) means that all stocks are given the same weight when calculating market and industry returns. Value-weighted (VW) means that stocks are given a weight based on their market capitalization when calculating market and industry returns. Stocks are allowed to switch industries multiple times. If stocks switch back to the industry they were previously in, this is treated as a seperate run of returns to prevent faulty linking of returns. Buy-and-hold stock returns are exactly matched to the benchmark on a monthly frequency.

Table 34: Percentage of stocks that outperform their benchmark over a decade horizon during 1970-2018

| | Non-stock benchmarks | | | Market benchmarks | | Industry benchmarks | |
|---|---|---|---|---|---|---|---|
| | % > 0 | % > T-bill | % > CPI | % > EW Mkt return | % > VW Mkt return | % > EW Ind return | % > VW Ind return |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 40 | 32 | 32 | 24 | 29 | 32 | 31 |
| Food Products | 64 | 54 | 54 | 40 | 45 | 34 | 39 |
| Candy & Soda | 70 | 61 | 59 | 40 | 51 | 35 | 51 |
| Beer & Liquor | 58 | 52 | 55 | 37 | 43 | 39 | 34 |
| Tobacco Products | 72 | 65 | 64 | 57 | 58 | 39 | 46 |
| Recreation | 37 | 30 | 31 | 21 | 25 | 30 | 30 |
| Entertainment | 41 | 36 | 37 | 29 | 30 | 31 | 28 |
| Printing & Publishing | 58 | 51 | 53 | 38 | 41 | 34 | 40 |
| Consumer Goods | 50 | 42 | 44 | 31 | 34 | 34 | 38 |
| Apparel | 50 | 40 | 42 | 31 | 35 | 39 | 38 |
| Healthcare | 47 | 41 | 42 | 33 | 34 | 32 | 38 |
| Medical Equipment | 46 | 41 | 42 | 32 | 33 | 31 | 30 |
| Pharmaceutical Products | 41 | 37 | 38 | 31 | 31 | 28 | 29 |
| Chemicals | 60 | 50 | 51 | 38 | 41 | 35 | 38 |
| Rubber and Plastic Products | 52 | 45 | 46 | 35 | 36 | 35 | 41 |
| Textiles | 46 | 38 | 38 | 27 | 32 | 37 | 36 |
| Construction Materials | 60 | 49 | 50 | 35 | 40 | 34 | 41 |
| Construction | 46 | 40 | 41 | 31 | 32 | 36 | 38 |
| Steel Works | 56 | 43 | 45 | 30 | 36 | 34 | 43 |
| Fabricated Products | 55 | 47 | 49 | 33 | 39 | 35 | 42 |
| Machinery | 57 | 48 | 50 | 35 | 39 | 33 | 43 |
| Electrical Equipment | 41 | 34 | 36 | 28 | 28 | 28 | 27 |
| Automobiles and Trucks | 52 | 43 | 46 | 32 | 32 | 37 | 38 |
| Aircraft | 65 | 56 | 56 | 40 | 45 | 32 | 37 |
| Shipbuilding, Railroad Equipment | 52 | 50 | 50 | 41 | 43 | 37 | 42 |
| Defense | 64 | 59 | 61 | 46 | 46 | 36 | 37 |
| Precious Metals | 29 | 24 | 25 | 21 | 21 | 28 | 28 |
| Mines | 38 | 28 | 30 | 25 | 25 | 32 | 30 |
| Coal | 51 | 44 | 44 | 32 | 33 | 43 | 40 |
| Petroleum and Natural Gas | 43 | 36 | 37 | 27 | 29 | 29 | 28 |
| Utilities | 85 | 73 | 73 | 45 | 55 | 38 | 59 |
| Communication | 50 | 46 | 46 | 38 | 38 | 33 | 37 |
| Personal Services | 44 | 39 | 40 | 32 | 32 | 35 | 38 |
| Business Services | 47 | 42 | 43 | 34 | 34 | 33 | 35 |
| Computers | 40 | 33 | 35 | 26 | 27 | 28 | 28 |
| Computer Software | 40 | 37 | 38 | 30 | 30 | 28 | 27 |
| Electronic Equipment | 47 | 40 | 42 | 31 | 33 | 28 | 36 |
| Measuring and Control Equipment | 53 | 43 | 46 | 32 | 35 | 26 | 34 |
| Business Supplies | 62 | 56 | 57 | 38 | 45 | 37 | 46 |
| Shipping Containers | 66 | 54 | 55 | 42 | 41 | 36 | 48 |
| Transportation | 53 | 46 | 48 | 34 | 37 | 39 | 37 |
| Wholesale | 47 | 41 | 42 | 33 | 33 | 34 | 34 |
| Retail | 50 | 43 | 44 | 33 | 36 | 36 | 36 |
| Restaurant, Hotels, Motels | 48 | 41 | 43 | 31 | 32 | 35 | 35 |
| Banking | 66 | 59 | 60 | 46 | 48 | 44 | 50 |
| Insurance | 66 | 57 | 58 | 41 | 46 | 38 | 45 |
| Real Estate | 56 | 50 | 52 | 37 | 34 | 40 | 41 |
| Trading | 64 | 55 | 57 | 45 | 43 | 40 | 40 |
| Other | 44 | 39 | 39 | 35 | 32 | 36 | 37 |
| Not available | 46 | 42 | 40 | 32 | 31 | 39 | 38 |
| Total | 51 | 43 | 45 | 33 | 35 | 35 | 38 |

All decade buy-and-hold common stock and REIT stock returns from January 1970 until December 2018 are benchmarked against zero, 30-day Treasury Bills, the Consumer Price Index (CPI), equal- and value-weighted market returns and equal- and value-weighted industry returns. Equal-weighted (EW) means that all stocks are given the same weight when calculating market and industry returns. Value-weighted (VW) means that stocks are given a weight based on their market capitalization when calculating market and industry returns. Stocks are allowed to switch industries multiple times. If stocks switch back to the industry they were previously in, this is treated as a seperate run of returns to prevent faulty linking of returns. Buy-and-hold stock returns are exactly matched to the benchmark on a monthly frequency.

Table 35: Percentage of stocks that outperform their benchmark over a lifetime horizon during 1970-2018

| | Non-stock benchmarks | | | Market benchmarks | | Industry benchmarks | |
|---|---|---|---|---|---|---|---|
| | % > 0 | % > T-bill | % > CPI | % > EW Mkt return | % > VW Mkt return | % > EW Ind return | % > VW Ind return |
| **Fama-French 49 Industries** | | | | | | | |
| Agriculture | 36 | 32 | 32 | 17 | 23 | 28 | 24 |
| Food Products | 61 | 52 | 53 | 39 | 43 | 31 | 38 |
| Candy & Soda | 68 | 62 | 61 | 39 | 50 | 28 | 53 |
| Beer & Liquor | 57 | 43 | 46 | 33 | 35 | 33 | 31 |
| Tobacco Products | 61 | 56 | 56 | 56 | 53 | 31 | 39 |
| Recreation | 33 | 24 | 25 | 17 | 19 | 25 | 23 |
| Entertainment | 35 | 30 | 31 | 22 | 24 | 26 | 23 |
| Printing & Publishing | 54 | 46 | 48 | 33 | 37 | 30 | 36 |
| Consumer Goods | 48 | 35 | 38 | 24 | 27 | 27 | 32 |
| Apparel | 43 | 35 | 37 | 25 | 29 | 32 | 30 |
| Healthcare | 45 | 38 | 40 | 29 | 30 | 29 | 33 |
| Medical Equipment | 43 | 37 | 38 | 26 | 27 | 27 | 23 |
| Pharmaceutical Products | 37 | 34 | 34 | 27 | 27 | 24 | 24 |
| Chemicals | 58 | 51 | 52 | 37 | 42 | 33 | 39 |
| Rubber and Plastic Products | 51 | 42 | 42 | 31 | 31 | 31 | 37 |
| Textiles | 50 | 39 | 38 | 24 | 30 | 36 | 34 |
| Construction Materials | 59 | 48 | 49 | 31 | 37 | 29 | 39 |
| Construction | 43 | 34 | 36 | 21 | 25 | 28 | 30 |
| Steel Works | 53 | 40 | 42 | 28 | 31 | 32 | 39 |
| Fabricated Products | 54 | 38 | 42 | 26 | 30 | 31 | 38 |
| Machinery | 53 | 44 | 46 | 30 | 33 | 28 | 39 |
| Electrical Equipment | 37 | 29 | 32 | 21 | 21 | 23 | 23 |
| Automobiles and Trucks | 49 | 40 | 41 | 24 | 25 | 31 | 30 |
| Aircraft | 64 | 55 | 54 | 38 | 45 | 26 | 34 |
| Shipbuilding, Railroad Equipment | 52 | 48 | 48 | 35 | 41 | 30 | 35 |
| Defense | 62 | 55 | 57 | 42 | 47 | 34 | 32 |
| Precious Metals | 26 | 21 | 23 | 16 | 14 | 21 | 23 |
| Mines | 34 | 27 | 29 | 23 | 24 | 32 | 29 |
| Coal | 52 | 46 | 48 | 31 | 33 | 44 | 41 |
| Petroleum and Natural Gas | 36 | 29 | 31 | 19 | 22 | 25 | 21 |
| Utilities | 79 | 73 | 75 | 44 | 56 | 35 | 57 |
| Communication | 45 | 40 | 41 | 31 | 31 | 31 | 32 |
| Personal Services | 42 | 36 | 36 | 26 | 28 | 31 | 34 |
| Business Services | 44 | 38 | 39 | 29 | 30 | 28 | 31 |
| Computers | 34 | 26 | 29 | 18 | 20 | 21 | 19 |
| Computer Software | 35 | 32 | 33 | 24 | 25 | 23 | 22 |
| Electronic Equipment | 42 | 35 | 36 | 24 | 25 | 22 | 30 |
| Measuring and Control Equipment | 50 | 38 | 41 | 25 | 28 | 20 | 28 |
| Business Supplies | 64 | 57 | 59 | 31 | 44 | 35 | 47 |
| Shipping Containers | 71 | 56 | 59 | 41 | 44 | 38 | 48 |
| Transportation | 47 | 42 | 43 | 31 | 32 | 36 | 34 |
| Wholesale | 43 | 37 | 38 | 28 | 28 | 29 | 28 |
| Retail | 46 | 38 | 39 | 28 | 32 | 32 | 32 |
| Restaurant, Hotels, Motels | 43 | 36 | 38 | 25 | 26 | 30 | 29 |
| Banking | 70 | 65 | 66 | 47 | 53 | 44 | 54 |
| Insurance | 67 | 57 | 58 | 38 | 44 | 35 | 44 |
| Real Estate | 55 | 48 | 50 | 34 | 33 | 37 | 37 |
| Trading | 69 | 61 | 64 | 48 | 45 | 39 | 41 |
| Other | 42 | 37 | 37 | 29 | 26 | 32 | 34 |
| Not available | 45 | 42 | 39 | 31 | 30 | 39 | 36 |
| Total | 48 | 40 | 41 | 29 | 30 | 31 | 33 |

All lifetime buy-and-hold common stock and REIT stock returns from January 1970 until December 2018 are benchmarked against zero, 30-day Treasury Bills, the Consumer Price Index (CPI), equal- and value-weighted market returns and equal- and value-weighted industry returns. Equal-weighted (EW) means that all stocks are given the same weight when calculating market and industry returns. Value-weighted (VW) means that stocks are given a weight based on their market capitalization when calculating market and industry returns. Stocks are allowed to switch industries multiple times. If stocks switch back to the industry they were previously in, this is treated as a seperate run of returns to prevent faulty linking of returns. Buy-and-hold stock returns are exactly matched to the benchmark on a monthly frequency.