Institutional ownership of sin and righteous stocks

Author:

Duco Kuiper

Student number: 412088

Supervisor:

Laurens Swinkels

Second assessor:

Esad Smajlbegovic

Master thesis Financial Economics

Erasmus University Rotterdam

Erasmus School of Economics

August 2018

Abstract

In this paper I examine and sin and righteous stocks and their relationship with institutional ownership using panel data for the years 1990 to 2016. I find that sin stocks have a significantly lower proportion of institutional ownership than non-sin stocks, when selecting sin stocks in the traditional way. Similarly, when selecting sin stock using different MSCI KLD ESG factors sin stocks also have a significantly lower institutional ownership than non-sin stocks. Furthermore, the newly introduced concept of righteous stock (opposite of sin stock) has a lower proportion of institutional ownership than non-righteous stock. This research also presents the behavior of sin stock institutional ownership over time and finds that the difference in proportion of institutional ownership between sin stocks and non-sin stocks has become larger over time. Furthermore, this paper also examines differences in institutional ownership for different sin industries and found that that there is no significant difference between the alcohol, tobacco and gambling industry regarding their relation to institutional ownership. Lastly, the hypothesis that more norm-constrained institutional investors hold a lower proportion of sin stocks than less norm-constrained investors cannot be accepted. The results regarding the differences between more and less norm-constrained institutional investors and their relationships with the proportion of sin and righteous stock ownership provide mixed evidence.

Keywords: Institutional ownership, Sin stocks, Sin industries, Righteous stocks and Social norm hypothesis[1].

[1] I would like to thank my thesis supervisor, Laurens Swinkels, for his supervision during the writing of my master thesis. His guidance, insightful suggestions, feedback and advise were very helpful and made it possible for me to write this thesis.

Table of contents

1. Introduction

Combining the topics of sin stocks, Environmental, Social and Governance (ESG) factors and institutional ownership is relatively new in the academic finance literature and is slightly gaining more popularity. According to Fauver & McDonald (2014), there has been an on-going debate in the academic and practitioner literature as to what impact social disapproval of smoking, gambling, and drink has on the equity values of firms that engage in the business of producing these goods (e.g., Fabozzi et al. (2008); Hong and Kacperczyk (2009) etc.). Previous research has shown that sin stocks and positive ESG factor stocks have a significantly different relationship with institutional ownership than other stocks (see for example Hong and Kacperczyk (2009) and Fernando, Sharfman and Uysal (2010). Even though there is increasing literature in the field of sin stocks, the majority is focused on sin stocks and returns and not so much on sin stocks and their ownership. The only published papers on the topic of sin stocks and ownership at this point are the papers of Hong and Kacperczyk (2009) and its Pacific-Basin sample replicate by Durand, Koh and Tan (2013). Given that these two articles are the only published articles on sin stocks and ownership it is important that more research is done on this topic in order to gain new insights. New insights are needed because much remains unclear about sin stocks and their ownership. For instance, we do not know whether there are differences between sin industries regarding their relationship to institutional ownership. Similarly it is unclear whether the relationship between sin stocks and institutional ownership has changed in recent years, and if so how the relationship has changed over time. With regards to ESG factors and institutional ownership, there is little consensus in the handful of papers. Fernando et al. (2010) find a negative relationship between ESG factors and institutional ownership. Graves & Waddock (1994) find a positive but insignificant relationship. Barnea & Rubin (2010) find no significant correlation and Coffey & Fryxell (1991) are the only authors that find a positive relationship between an ESG factor and institutional ownership. From the lack of concensus it becomes clear that new research on these topics is required, because it remains vague what the real relationship between ESG factors and institutional onwership is. Acquiring more knowledge on these relationships between sin stocks, positive ESG stocks and institutional ownership is valuable because it contains information on how institutional investors treat firms that act socially, ethically or economically wrong or right according to society. This information is potentially useful for for example policy makers who want to promote or disencourage certain behaviour of firms. More specifically, if for example the data entails that inverstors shun away from socially well performing firms leading to a shortage of funds for these firms, then policy makers may decide that subsidies are necessary. Additionally, the topics of ESG factors and socially responsible investing (SRI) have become increasingly popular among practitioners and academics in recent years (see Skypala (2017) and Capelle-Blancard & Monjon (2012)). One manifestation of this is the recent launch of the

MSCI New Factor ESG Target Indexes in September 2017. Previous papers on these topics are potentially outdated due to the shift in popularity of these topics. This development is another reason why more recent research is required.

The definition of sin stocks has shown to differentiate in previous literature because the definition of that what constitutes a controversial industry is itself controversial. I utilize two methods of selecting sin stocks. For the first method, I will adhere to the definition of Hong and Kacperczyk (2009) and Salaber (2007) to select sin stocks. This means that the firms involved with alcohol, tobacco and gambling will be characterized as sin stock firms. This method of selecting sin stocks is refered to as the traditional method throughout this paper. Even though the sex industry and the weapons industry can be seen as sinful, I will not include these industries in my main research. This is in line with previous literature on sin stocks (see for example Hong and Kacperczyk (2009) and Salaber (2007)). Salaber (2007) states that the defense or weapons industry is not as homogenous as the other three sin industries. Firstly, because the different firms in the weapons industry have diverse activities. Secondly, because alcohol, tobacco and gambling are considered sinful for religious reasons. Thirdly, because alcohol, tobacco and gambling have an addicitve aspect wih implied external health care cost. Lastly, because alcohol, tobacco and gambling are subject to excise taxes[2]. Hence, the weapons industry is not included as a sin industry. Previous literature of for example Hong and Kacperczyk (2009) has shown that the weapons industry does not significantly affect sin stock and ownership results. However, at the end of this paper I will execute a robustness check including the weapons industry to show whether the results change. Furthermore, it is unlikely that including the sex industry will affect my results because only very few publicly traded companies exist. Since the number of companies is small, both in absolute terms and in relation to the total number of firms, the results will not be affected in a significant way. The second group of sin stocks is selected using MSCI KLD ESG factors. This group differs from the previous sin group because the definition of 'sin' is not limited to the alcohol, tobacco or gambling industry. Firms that negatively stand out on ESG factors are selected as sin stocks. These factors include human rights, child labor, waste management and climate change. I deliberately use other ESG factors than the also available MSCI KLD ESG factors involvement in alcohol, tobacco and gambling. The reason for this is that I am interested in the effects of the social norm hypothesis when sin stocks are not selected using the traditional way (using alcohol, tobacco or gambling). The group of stocks that are characterized as righteous stock are also selected using MSCI KLD ESG factors. The term righteous stock denotes the opposite of sin stocks and describes firms that

[2] For more details on the reasons why the weapons industry is not included in the sin definition see Salaber (2007).

positively stand out on the ESG factors human rights, waste management and climate change. The methods to select sin and righteous stock will be descibed in greater detail in chapter 3.2.

In this paper I examine the relationship between institutional ownership and sin and righteous stocks using panel data for the years 1990 to 2016. The main question is therefore; *what is the relationship between institutional ownership and sin and righteous stocks?* I construct four main hypotheses around this topic that all research subcategories of this question. These subcategories include time differences, industry differences and shareholder type differences. The theory behind the first hypothesis is the social norm hypothesis of Hong and Kacperczyk (2009). Institutional investors are more prone to public scrutiny in comparison to individual investors and their investing has a more public character. This means that they invest relatively less in stocks that could potentially harm their reputation, because a loss in reputation could lead to further negative consequences such as diminishing future cash inflows. I expect that the sensitivity to reputation for institutional investors works two ways. Investing in sin stocks decreases the reputation of investors and investing in righteous stocks increases the reputation of investors, whereby an increase in reputation has positive consequences such as increasing future cash inflows. In line with previous research such as Hong and Kacperczyk (2009) I hypothesize the following:

H1A: The proportion of institutional ownership of sin stocks is lower than the proportion of institutional ownership of non-sin stocks.

H1B: The proportion of institutional ownership of righteous stock is higher than the proportion of institutional ownership of non-righteous stock.

The first hypothesis describes the relationship between institutional ownership and sin stocks in general and without any distinction between time periods. The next hypothesis adds value to this paper by analyzing the relationship between institutional ownership and sin stocks over time. There are two main reasons why I expect changes in the institutional ownership of sin stocks. The first one involves average institutional ownership. The percentage institutional ownership of stocks has changed over time. The proportion of equity holdings of institutional investors amounted to five percent between 1900 and 1945 (Blume & Keim, 2014). However, Blume & Kleim (2014) find that this percentage increased after the second world war. Their research shows that 34 percent of equity is held by institional investors in 1980 and 67 percent is held by 2010. The second reason involves the popularity and practice of SRI over time. Capelle-Blancard & Monjon (2012) report a clear upward trend in the popularity of SRI over the years 1982 to 2009 for both public literature and academic literature. See appendix graph 2 for a visual overview of the SRI popularity findings of Capelle-Blancard & Monjon (2012). Due to increased popularity of SRI over time I expect a change in time-series data

on sin stock institutional ownership. On average I expect that institutional ownership has increased over time, however I expect that this increase is of a smaller order for sin stocks. I expect that sin stocks get shunned more by institutional investors in later years than in earlier years of the sample. I construct the following hypothesis:

H2: The difference in proportion of institutional ownership between sin stocks and non-sin stocks becomes larger over time.

The third hypothesis contributes to this paper by reviewing the difference in relationship between sin stock institutional ownership and different sin industries. Due to different societal viewpoints towards the different sin industries and different exclusion lists of institutional investors I expect that the different sin industries will have significantly different coefficients. This means that institutional investors shun different sin industries more or less extensively. For example, smoking and the tobacco industry receive more negative attention from governments, in the form of anti-marketing campaigns, and from the medical industry, in the form of lawsuits, than the gambling industry. Additionally, the UN Global Compact excluded tobacco companies from participating in the UN Global Compact in 2017. The alcohol and gambling industry were not excluded, meaning that the industries are viewed differently. The hypothesis is therefore as follows:

H3: The different sin industries have a significantly different relationship with the proportion of institutional ownership.

Thus far, all instiutional investors are treated the same. The next hypotheses will differentiate between different institutional investor classes. The difference is made between more and less norm-constrained institutional investors. Being more norm-constrained means that investors are more constained by the norms that society put on them. More norm-constrained investors are therefore more affected by the social norm than less norm-constrained investors, leading to an expected difference in behaviour. Similar to the research of Hong & Kacperczyk (2009) I expect more norm-constrained institutional investors to hold relatively less sin stocks in comparison to less norm-constrained institutional investors. This means that institutional investors such as pension funds have a different relationship with sin stocks than for example hedge funds. Pension funds, university endowment funds, religious organisations, banks and insurance companies are considered more norm-constrained institutional investors and hedge funds and mutual funds are considered less norm-constrained institutional investors. I also expect that more norm-constrained institutional investors have a different relationship with righteous stock than less norm-constrained institutional investors. In short, the hyptheses follow the social norm hypothesis. The former leads to the following hypotheses:

H4A: The proportion of sin stock ownership of more norm-constrained institutional investors is relatively lower than the proportion of sin stock ownership of less norm-constrained institutional investors.

H4B: The proportion of righteous stock ownership of more norm-constrained institutional investors is relatively higher than the proportion of righteous stock ownership of less norm-constrained institutional investors.

With regards to the hypotheses I find the following. The results show that when selecting sin stocks in the traditional way, sin stocks have a significantly lower proportion of institutional ownership than non-sin stocks. Similarly, when selecting sin stock using different MSCI KLD ESG factors sin stocks also have a lower proportion of institutional ownership than non-sin stocks. Furthermore, righteous stocks have a lower proportion of institutional ownership than non-righteous stocks. This research also presents the behavior of sin stock institutional ownership over time and finds that the relationship between sin stocks and institutional ownership has become more negative over time. This means that the difference in proportion of institutional ownership between sin stocks and non-sin stocks has become larger over time. Furthermore, this paper also examines differences in institutional ownership for different sin industries and finds that there is no significant difference between the three sin industries and their relationship to institutional ownership. Lastly, the hypothesis that more norm-constrained institutional investors hold a lower proportion of sin stocks than less norm-constrained investors cannot be accepted. The results regarding the differences between more and less norm-constrained institutional investors and their relationships to the proportion of sin and righteous stock ownership provide mixed evidence.

This paper contributes to the existing literature in the following ways. First of all this paper introduced a new way of selecting sin stocks and examined the sin stock and institutional ownership relationship using this new method. This adds value to the literature by providing more evidence on the robustness of the social norm hypothesis. Second of all, this paper introduced a new term: righteous stocks, and examined the relationship between righteous stock and institutional ownership. Evidence on the relationship between institutional ownership and sin and righteous contributes to the literature because it contains information on how institutional investors treat firms that act socially, ethically or economically wrong or right according to the norms of society. This paper shows that institutional investors in general 'punish' both sin activities and righteous activities by lower institutional ownership. This information is potentially useful for policy makers, owners of these firms or even new firms who are considering new business activities that contain aspects of sin or righteous factors. Third of all, this paper presents the behavior of the relationship between sin stocks and institutional

ownership over time for recent years, which provides more information on how institutional investors treat sin stocks. Fourth of all, this paper adds value by zooming in on the relationship of sin stocks and institutional ownership by examining different sin stock industries and their individual relationship to institutional ownership. This has not been done before according to the best of my knowledge and therefore adds to the completeness of the state of the academic literature (contribution for academics). The reason why this contributes for practitioners is because it provides information on whether investors distinguish between different sin industries. This paper shows that there is no significant difference between sin industries regarding their relation to institutional ownership. For instance, this is then valuable information for a sin industry firm looking to invest in another sin industry firm because the firm knows that is not likely to incur an additional institutional ownership penalty for this investment[3]. Finally, this paper contributes by testing the social norm hypothesis with more recent data, thereby testing whether social norms still have an effect on markets in recent years.

The remainder of this paper is organized as follows. Section 2 covers the theoretical framework. Section 3 describes the data and includes summary statistics, data cleaning methods and variable descriptions. Section 4 presents the methodology. Section 5 consists of the results followed by section 6 which includes a series of robustness checks. Section 7 covers the discussion of the results and section 8 consists of the conclusion. Finally, section 9 and 10 cover the reference list and the appendix respectively.

2. Theoretical framework

The theoretical framework is build up as follows. Firstly, I zoom out of sin and righteous stocks and discuss the topic of corporate social responsibility (CSR) and socially responsible investing (SRI). This helps to understand why the discussion around sin stocks and righteous stocks is important in the first place. After introducing CSR and SRI I focus on CSR, SRI and (institutional) ownership. Thirdly, I focus on sin stocks returns and on sin stocks ownership specifically and discuss the current literature regarding this topic.

2.1 Introduction to CSR and SRI

The topic of CSR emerged more extensively in the second half of the 20th century, resulting from a growing interest in the impact of companies' activities on larger society (see Carroll, 1999). Authors such as Brigham, Gapenski, and Ehrhardt (1999, p. 3) asked questions such as: ''How do we balance

[3] This reasoning only holds with regards to this research. Incurring an institutional ownership penalty remains possible for several other reasons.

social concerns against the need to create value for our shareholders?'' CSR can bring about positive financial performance so neglecting CSR could result in losses or missed opportunities. On the other hand, managers who focus extensively on creating social value might neglect the primary reason for their company's existence. As an illustration as to how to CSR can affect financial performance; Goss & Roberts (2011) find that firms with social responsibility concerns pay between 7 and 18 basis points more on bank debt than firms that are more responsible. Nowadays the practical implications of CSR receive attention from all angles. Research suggests that corporate social responsibility represents a differentiating factor that may be used successfully by firms to distinguish themselves within their industries (Drumwright, 1994).

Very closely related to the topic of CSR is the topic of SRI. I follow authors such as Sethi (2005) and define SRI according to the Social Investments Forum, the industry association of the SRI in the United States. Their definition is as follows*:*

Socially responsible investing (SRI) is investing in companies that meet certain baseline standards of social and environmental responsibility; actively engaging those companies to become better, more responsible corporate citizens; and dedicating a portion of assets to community economic development.

SRI has origins in ancient Jewish, Christian and Islamic traditions and is not new to the world[4]. The scale of SRI has changed however. Already in 2003 Krumsick noted, ''socially responsible investing ... is more wide-spread than ever, in the U.S. as well as Europe and Asia.'' The magnitude of socially responsible investing has grown to where one out of every eight dollars placed in the hands of professional money managers in the United States is dedicated to such investment opportunities (Laufer, 2003). SRI has changed from an activity carried out by a small number of specialist retail investment funds (in the form of unit trusts and mutual funds), probably of negligible or minor economic importance, into an investment philosophy adopted by a growing proportion of large investment institutions, i.e. large pension funds and insurance companies (Sparkes & Cowton, 2004). Sparkes and Cowton (2004) argue that this shift in SRI from margin to mainstream could play a crucial role in obliging or influencing quoted companies to address CSR issues. CSR factors are used by (institutional) investors such as pension funds to participate in SRI. Socially responsible investors invest using filter screens whereby either firms that perform poorly on CSR measures are excluded or firms that perform well on CSR measures are included in investment portfolios.

[4] For an extensive description of the history of SRI see Renneboog, Ter Horst, & Zhang (2008).

2.2 CSR, SRI and stock ownership

Several researchers have researched the relationship between corporate social responsibility, socially responsible investing and ownership. Fernando et al. (2010) find that both positive (“green”) and negative (“toxic”) environmental performers have a larger number of shareholders but lower institutional ownership. The social norm hypothesis can explain why investors shun away from negative environmental stocks but it does not explain the relationship between positive environmental performers and institutional ownership that is found in Fernando et al. (2010). They also find a difference in holdings for the different types of institutional investor. All institutional types except other institutions (including universities, pension plans and employee stock ownership plans) hold significantly lower fractions of the shares of green firms. In contrast, only other institutions hold a significantly smaller percentage of shares of toxic firms. The authors pose the possible explanation that institutions may shun green stocks due to a belief that corporate investment in positive environmental performance detracts from shareholder value. Graves & Waddock (1994) find a significant positive relationship between social performance and the number of institutions holding the shares of a company and a positive but insignificant relationship between social performance and the percentage of shares held by institutions. They conclude by stating that improving a company's corporate social performance invokes no penalty in institutional ownership. Barnea & Rubin (2010) report that on average, insiders’ ownership and leverage are negatively related to the firm’s social rating, while institutional ownership is uncorrelated with it. Some authors also focus on very specific CSR measures. Coffey & Fryxell (1991) for example find a positive relationship between the amount of institutional ownership of corporate stock and a company's social responsiveness as measured by the representation of women on its board of directors and they found no statistically significant relationship with social responsibility as measured by charitable giving.

When comparing and contrasting these findings distinct contrasts and similarities are found. The first two papers find similar results regarding the effect of ESG factors on the number of institutional shareholders. Both papers find a positive relationship between an ESG factor (environmental for Fernando et al. (2010) and social for Graves & Waddock (1994)) and the number of institutional shareholders. However, contradictory results are found regarding the percentage of institutional ownership. In Fernando et al. (2010) positive (“green”) and negative (“toxic”) environmental performers have a lower proportion of institutional ownership and in Graves & Waddock (1994) social performancers have a positive but insignificant relationship with the percentage of shares held by institutions. Similarly as to Graves & Waddock (1994), Barnea & Rubin (2010) find no significant correlation between a firm’s social rating and institutional ownership. Coffey & Fryxell (1991) are the

only authors that find a positive relationship between an ESG factor and the amount of institutional ownership.

The differences between different institutions, regarding investing in sin/toxic or green stock, could be explained by different investment priorities or characteristics. Derwall, Koedijk, & Ter Horst (2011) differentiate between two types of investors; the values-driven and profit-seeking social investors. They describe that values-driven social investors' motivation behind the investor's decision to shun controversial stocks varies between purely emotional (see for example Beal and Goyen (1998) and Statman (2005)) and societal driven (see for example Hong and Kacperczyk (2009). The profit-seeking social investor is merely focussed on economic profit. Bollen (2007) argues that investors may have a multi-attribute utility function that is not only based on the standard risk-reward optimization but also incorporates a set of personal and societal values. Bollen (2007) describes that these socially responsible investors consume the socially responsible attribute of the investments.

Another difference between different institutional investors is the behaviour after purchasing a stock. Some investors attempt to actively influence firm's corporate social responisibility behaviour. Pension fund equity is positively related to both a people (women and minorities, community, and employee relations) and a product quality (product and environment) dimension of CSP, but mutual and investment bank funds exhibit no direct relationship with CSP (Johnson & Greening, 1999). Neubaum & Zahra (2006) show that long-term institutional ownership is positively associated with CSP and that the frequency and coordination of activism interact with long-term institutional holdings to positively affect CSP 3 years later.

2.3 Sin stock returns and sin stock ownership

Valuable research has been done in the field of sin stocks, however much more has to be researched. Most of the existing research focusses on sin stocks and returns and not so much on ownership structures. The literature on sin stock returns is mixed; the returns on sin stocks are not always significantly better or worse than comparable non-sin stocks[5]. Hong and Kacperczyk (2009) find that sin stocks have higher expected returns than comparable stocks, consistent with them being neglected by norm-constrained investors and facing greater litigation risk heightened by social norms. Salaber (2007) narrows this higher expected return down to legal and religious factors. His results show that sin stocks have higher risk-adjusted returns when they are located in a country with high excise taxation; and sin stocks outperform other stocks when the litigation risk is higher. Salaber (2007) also finds that Protestants are more "sin averse" than Catholics, and require a significant premium on sin stocks. Salaber's findings suggest that sin stock returns depend on both legal and religious

[5] See Derwall, Koedijk, & Ter Horst (2011) table 1 for an overview of sin stock returns literature.

environments of each country. Lobe and Walkshäusl (2016) find no compelling evidence that sin stocks, or socially responsible stocks outperform or underperform. Statman and Glushkov (2009) find evidence for the "no effect" hypothesis, whereby the expected returns of socially responsible stocks are approximately equal to the expected returns of conventional stocks. Last, Durand, Koh and Tan (2013) find that sin stocks generate negative risk-adjusted returns in each of the seven Pacific-Basin markets they analyze. The recent article of Blitz and Fabozzi (2017) claims that the abnormally high raw returns can be fully explained by the new quality factors of Fama and French (2015), profitability and investment. The anomaly of high returns for sin stocks is therefore resolved according to them.

Hong and Kacperczyk (2009) are the main contributors to literature on sin stocks and ownership. They find that that sin stocks are less held by norm-constrained institutions such as pension plans as compared to mutual or hedge funds during the period of 1980-1997. The social norm hypothesis is particularly important for this research. Akerlof (1980) describes that social customs which are disadvantageous to the individual may nevertheless persist without erosion, if individuals are sanctioned by loss of reputation for disobedience of the custom[6]. The social norm hypothesis entails that investors are subject to the social norm and that not complying with this norm is punished by a loss of reputation which can have negative effects in the form of for example diminished future capital inflow. Different types of investors have a different sensitivity to the social norm and a loss of reputation. The loss of reputation that can occur for (institutional) investors when investing in sin stock can have a significant effect on sin stock ownership (see for example Hong and Kacperczyk (2009))[7]. Hong and Kacperczyk (2009) find that sin stock comparables (defined as those with similar Fama-French (1997) industry groupings as our sin stocks) have on average about 22% of their shares held by institutions. In contrast, sin stocks have about 19% of their shares held by institutions, which is approximately a 14 percent lower institutional ownership ratio than their comparables (Hong & Kacperczyk, 2009).

The research of Hong and Kacperczyk (2009) has been repeated by Durand, Koh and Tan (2013) for a Pacific-Basin sample including Australia, New Zealand, Japan, South Korea, India, Malaysia and Singapore. In Australia and New Zealand, the countries which are culturally closest to the US, they find that institutions are less likely to hold sin stocks[8]; such a finding is consistent with Hong and Kacperczyk's (2009) social norms hypothesis. In Japan and South Korea, they find that substantial

[6] Multiple definitions of social customs or norms exist in the literature. I follow Akerlof (1980) who defines a social norm or custom as an act whose utility to the agent performing it depends in some way on the beliefs or actions of other members of the community

[7] The definition of reputation has been described as 'a perceptual representation of a company's past actions and future prospects that describe the firm's overall appeal to all its key constituents when compared to other leading rivals' (Fombrun, 1995, p. 72).

[8] Durand, Koh and Tan (2013) use substantial holdings to proxy for institutional holdings.

shareholders are more likely to hold sin stocks. In the other markets they find no difference in substantial shareholdings of sin stocks. The results are therefore only partly similar to the research Hong and Kacperczyk (2009).

As explained in Fauver & McDonald (2014) past research has assumed that a sin stock in one country is automatically a sin stock in another country (see for example, Hong and Kacperczyk (2009), Statman and Glushkov (2009), Fabozzi et al. (2008)). More recent research has shown that this assumption is not trustworthy because of differences in societal views towards sin stocks (see for example Roca and Wong (2010), Derwall et al. (2011), and Durand et al. (2013)). As mentioned in Fauver & McDonald (2014); Stulz and Williamson (2003), Heinrichs et al. (2006), and Kumar et al. (2011)) doubt this assumption. This means that an alcohol stock could be seen as a sin stock in one country and could be seen as a regular stock in another. Activities deemed acceptable or appropriate by one constituency may be viewed as undesirable or immoral by another (Robin and Reidenbach, 1987). This stream of literature could also be used to argue that the different sin stock industries could be viewed differently within a society. For example, tobacco might be seen as more sinful by the majority of society than alcohol. Other authors such as Fabozzi, Ma, & Oliphant (2008) reported results on differences between different sin industries. However, this has only been done for sin stocks and return and not for sin stocks and ownership. Hence, there is a gap in current literature on the differences between the different sin industries and institutional ownership.

The importance to view sin stocks industries individually next to the sin industry as a whole can also be seen from comparing exclusions lists from multiple investors. Different investors use different factors for their exclusion lists which could lead to different outcomes per sin stock industry. As an example, sizeable funds and investors such as ABP/APG, PGGM, PGGM and Robeco exclude tobacco stocks but do not exclude gambling stocks. This can possibly result in relatively lower institutional ownership of tobacco stocks in comparison to gambling stocks because of investing exclusion. In conclusion, the heterogeneity among investors regarding their exlusion lists is a potential driver for finding different relationships between the different sin industries and institutional ownership.

3. Data description

The data section of this paper consists of multiple subsections. The first section describes the data and the sample. The second section elaborates on the sin stock and righteous stock selection process and the description of the MSCI KLD ESG factors. The third section consists of the description and construction of all remaining variables. The last section includes the summary statistics of the data.

3.1 Data and sample selection

The data that is used for this research paper is collected from three main databases. Firstly, data on institutional ownership and (sin) stocks is retrieved from Thomson Reuters 13f fillings. The two main variables that are taken from this database are Total institutional ownership as a percentage of shares outstanding and Shares held per shareholder type. This database initially provides quarterly data from the 31st of March 1990 to the 31st of December 2016, but this is modified to annual data by removing the first three quarters for the years whereby the last quarter (31st Dec) is available. If the last quarter is not available then the data of the third, second or first quarter is used (in that order). The reason for modifying this data is to match the CRSP/Compustat data, which is annual data. This eventually results in panel data of American stocks and their institutional ownership per year.

The second database is CRSP/Compustat which is a merged database consisting of CRSP and Compustat data. The combined database of CRSP and Compustat is used to collect data on stock returns, stock characteristics and firm fundamental accounting values. The main variables provided by this database are SIC codes, NAICS codes, Liquidation value of common equity, NASDAQ and S&P500 constituents, Beta, Total market value, Stock price, Average monthly return and Standard deviation of daily returns. These variables originate from different sub datasets within CRSP/Compustat. The initial data consists of/ or is modified to annual data per stock, listed in the United States of America, for the years 1990 to 2016.

The third and final database, MSCI KLD ESG Stats, adds value in the form of data on ESG factors which leads to the selection of sin stocks and righteous stocks. MSCI KLD ESG Stats is a database that includes data on environmental, social and governmental factors of firms. Graves & Waddock (1994) and Fernando et al. (2010) argue that the KLD data is the best single source of social and environmental performance measures because of the expertise and objectivity of the analysts who assign the KLD ratings and the wide range of attributes across which these ratings are assigned. Other authors such as Statman & Glushkov (2009) and somewhat more recently Zyglidopoulos, Georgiadis, Carroll, & Siegel (2012) have also used the MSCI KLD ESG database. This database provides firm level data on the social factors human rights and child labor and on the environmental factors waste management and climate change. These variables are further explained in the next section. The MSCI KLD ESG stats database reports back until 1991, however CUSIP data is only available from 1995. Having CUSIP data is essential for matching the MSCI KLD ESG Stats dataset to the other datasets. Therefore research using this data covers the time span 1995 to 2016. The individual ESG variables also have different data availabilities. The individual variables have data for the following years: child labor 2012 to 2016, positive human rights 1995 to 2016, negative human rights 2012 to 2016, positive climate change 1995

to 2016, negative climate change 1999 to 2016, positive waste management 1991 to 2016 and negative waste management 1991 to 2016.

Following current literature I exclude stocks from my data that have a one-digit SIC code of 6. This SIC code stands for the financial service industry. The reasoning behind excluding this industry is that financial service companies are under governmental legislative constraints and because of the typically high leverage described by Fama & French (1992). This leads them to not be representative of the sample.

3.2 Sin stock and righteous stock selection process

I follow other authors such as Hong & Kacperczyk (2009) in their sin stock identification method for my first sample of sin stocks. In order to identify sin stocks and label them as such I use the Fama & French (1997) stocks classification which uses SIC codes to create 48 different industry classifications. Industry group 4 (beer or alcohol) and industry group 5 (smoke or tobacco) are classified as sin industries and the stocks that fall within these industries are marked as sin stocks. Stocks with SIC codes 2100-2199 belong to the tobacco group, and those with SIC codes 2080-2085 belong to the beer group. The identification of gambling sin stocks requires extra information since the SIC codes do not make distinctions between hotel stocks, entertainment stocks and gambling stocks. The NAICS classification does have a separate industry class for gambling stocks. Therefore, the NAICS codes 7132, 71312, 713210, 71329, 713290, 72112 and 721120 are used to identify gambling stocks, which will also be marked as sin stocks.

The second method to select sin stocks utilizes the MSCI KLD ESG factors. Stocks that are denoted as ‘1’ on the negative MSCI ESG factors human rights, child labor, waste management and climate change are denoted as sin stocks. The meaning of being denoted as ‘1’ is described in table 1. Stocks that score a ‘1’ on one of the factors are denoted as sin stocks for that year. The selection of righteous stocks happens in a similar way. Stocks that are denoted as ‘1’ on the positive counterpart of MSCI KLD ESG factors human rights, waste management and climate change are denoted as righteous stocks. Once again, when in a certain year a stock scores a ‘1’ on at least one of the positive factors, then the stock is denoted as a righteous stock.

Table 1
Description of MSCI KLD ESG factors
(+) stands for a positive factor and (-) stands for a negative factor.

Factor	Description[9]
Child labor (-)	The severity of child labor controversies in a firm's supply chain. Factors affecting this evaluation include, but are not limited to, a history of involvement in child labor-related legal cases, widespread or egregious instances of child labor in the firm's supply chain, resistance to improved practices, and criticism by NGOs and/or other third-party observers.
Waste management (+)	This indicator assesses how companies manage their risk of incurring liabilities associated with pollution, contamination, and the emission of toxic and carcinogenic substances. Companies that have strong programs and performance in reducing toxic emissions score higher.
Waste management (-)	The severity of controversies related to a firm's non-GHG emissions. Factors affecting this evaluation include, but are not limited to, a history of involvement in land or air emissions-related legal cases, widespread or egregious impacts due to hazardous emissions, resistance to improved practices, and criticism by NGOs and/or other third-party observers. (formerly known as toxic spills & releases)
Climate change (+)	The positive side assesses how companies manage the risks of increased costs linked to carbon pricing or regulatory caps. Companies that proactively invest in low-carbon technologies and increase the carbon efficiency of their facilities or products score higher. Management metrics include efforts to reduce exposure through comprehensive carbon policies and implementation mechanisms, including carbon reduction targets, production process improvements, installation emissions capture equipment, and/or switching to cleaner energy sources.
Climate change (-)	The severity of controversies related to a firm's climate change and energy-related policies and initiatives. Factors affecting this evaluation include, but are not limited to, a history of involvement in GHG-related legal cases, widespread or egregious impacts due to corporate GHG emissions, resistance to improved practices, and criticism by NGOs and/or other third-party observers.
Human right policies & initiatives (+)	The human right policies & initiatives indicator identifies companies that have undertaken exceptional human rights initiatives, including outstanding transparency or disclosure on human rights issues, or has otherwise shown industry leadership on human rights issues not covered by other MSCI human rights ratings.
Human rights violations (-)	The severity of controversies related to the impact of a firm's operations on human rights. Factors affecting this evaluation include, but are not limited to, a history of involvement in human rights-related legal cases, widespread or egregious complicity in killings, physical abuse, or violation of other rights, resistance to improved practices, and criticism by NGOs and/or other third-party observers.

3.3 Variable description and construction

The retrieved observations from the different databases are merged using CUSIP and Year to obtain a dataset with one observation for every variable per stock per year. All duplicates are removed based

[9] All ESG factor variable descriptions are taken from the MSCI ESG KLD methodology manual. For more details on the MSCI ESG factors see the MSCI ESG KLD methodology manual.

on CUSIP and Year for all three datasets prior to merging the different databases. The first two variables are defined as follows. Percentage institutional ownership is defined as shares owned of stock i by institutional investors for year t divided by the total number of shares outstanding of stock i for year t. This delivers a percentage of institutional ownership per stock for all available years. Observations whereby percentage of institutional ownership exceeds 100% are winsorized to 100% because having more than 100% ownership is theoretically and practically not possible. The variable shareholder type contains 5 numerical values: 1=bank, 2=insurance company, 3=investment companies and their managers, 4=independent investment advisor, 5=all others. All others includes pension plans, university endowment funds and religious organizations. The dependent variable Percentage institutional ownership is constructed two more times in a different way in order to answer the last hypothesis. For every stock i the total shares owned by shareholder types 1, 2 and 5 are summed up for year t and divided by total number of shares outstanding for year t. The same method is applied for less norm-constrained investors, which means that only shareholder types 3 and 4 are included. The shareholder type data suffers from measurement errors. Gompers and Metrick (2001) document the evolution of institutional investors and show that the database Thomson Reuters improperly classifies institutions in the first four categories into group 5, with the vast majority of investment advisors being mistakenly allocated to "all others" group. This classification is wrong starting from the last quarter of 1997 and beyond[10]. This means that institutional investor type 4 was added to investor type 5. Meaning that less norm-constrained institutional investors are added to the list of more norm-constrained institutional investors. In order to tackle this issue I exclude shareholder type 5 from the data from the last quarter of 1997 and after.

The sin and righteous stock identification is followed by dummy variable construction. Five sin dummy variables are created. The first one (Sindum) is constructed twice using two separate methods. The first one simply denotes a stock as a sin stock or non-sin stock based on the previously mentioned industry categories (alcohol, tobacco and gambling) and will be referred to as Sindum1. The variable takes on the value one if the stock is a sin stock and zero if the stock is not a sin stock. The Sindum variable is also constructed using the factors from MSCI KLD ESG database whereby a stock is denoted as a sin stock using the process described in section 3.2. This sin stock dummy will be referred to as Sindum2. Similarly, the variable takes on the value one if the stock is a sin stock and zero if the stock is not a sin stock. The following dummy is Rightdum which is short for righteous stock dummy. The construction of this variable also utilizes MSCI KLD ESG data. Stocks are denoted as righteous when they score a '1' for one of the positive ESG factors. This is also described in section 3.2. The variable takes on the value one if the stock is a righteous stock and zero if the stock is not a righteous stock.

[10] For a full explanation see Gompers and Metrick (2001).

The next three dummy variables are variables for the separate sin stock industries. Alcdum is a dummy variable that is equal to one if the stock is an alcohol sin stock (SIC codes 2080-2085) and zero otherwise. Gamdum is a dummy variable that is equal to one if the stock is a gambling sin stock (NAICS codes 7132, 71312, 713210, 71329, 713290, 72112 and 721120) and zero otherwise. Tobdum is a dummy variable that is equal to one if the stock is a tobacco sin stock (SIC codes 2100-2199) and zero otherwise.

The control variables for the regressions are mainly measured in a similar manner as in Hong & Kacperczyk (2009) and are measured or constructed as follows. The first two control variables are variables to control for institutional ownership of certain industries. I construct two industry dummies. The first one is ONEDIGDUM which is a dummy variable that takes on the value one if the one digit sic code coincides with the one digit sic code of the alcohol, tobacco or gambling industry (codes 2 and 7), and zero otherwise. The second industry dummy is a more specific measure for industry group. I take the Fama and French (1997) 49 industry categories and design a dummy variable FFDUM which takes on the value one if it falls in one of the following categories: 2 (food), 3 (soda), 4 (beer), 5 (smoke), 7 (fun), 43 (meals), and 49 (gambling). If the industry category does not fall into one of the previously mentioned then FFDUM takes on the value zero. These dummies control for institutional ownership of comparable consumer stocks. The next control variable is LOGSIZE which is the measure for size and is calculated by taking the natural logarithm of market capitalization. Market capitalization is calculated as end of year share price on year t for firm i times shares outstanding on end of year t for firm i. LOGMB is the natural logarithm of the market to book value of firm i for year t and is measured as the end of year market capitalization divided by the end of year book value. The variable SP500 is a dummy variable that takes on the value one for year t if stock i is listed on the S&P 500 index for year t and zero otherwise. The rather similar variable NASD is a dummy variable that takes on the value one for year t if stock i is listed on the NASDAQ for year t and zero otherwise. STD is the standard deviation of daily returns for firm i during year t, whereby returns are the simple raw returns. BETA is the year-end beta reported by Compustat in year t for firm i. PRINV is the inverse of the firm i share price at the end of year t. RET is the average monthly return on stock i during year t. Prior to empirically testing the hypotheses I remove all observations that have a missing value for any of the dependent or control variables.

3.4 Summary statistics

The final sample consists of 81,662 observations and 10,855 different firms. The summary statistics in this paper are on some grounds similar and on others different from the summary statisics in Hong & Kacperczyk (2009). The average percentage of institutional ownership of my more recent dataset (1990-2016) is 0.430 which is larger than the older dataset (1980-2003) of Hong & Kacperczyk (2009)

who report an average of 0.27 (see table 2). Graph 1 shows that average institutional ownership has fluctuated over time. Graph 1 shows a clear upward trend for average institutional ownership for the years 1990 to 2007. The relation after 2007 is rather stable at first and slightly decreasing afterwards. Moreover, the mean of Logsize is 19.631 and the mean of Logmb is 0.771 with standard deviations of respectively 2.270 and 0.962. These results are comparable to the results of Hong & Kacperczyk (2009). However, the sample of this paper contains larger firms because logsize is larger (19.631 compared to 11.16). The reason for this could be changing firm size over time. Average firm size in the sample of this paper is smaller in the first few years and larger in the later years. Poschke (2011) finds that average firm size has increased over time in the U. S. so this offers an explanation for the observed summary statistics. The means and standard deviations of the variables STD (0.036 and 0.024) and RET (0.013 and 0.052) in this paper are similar to the reported values in Hong & Kacperczyk (2009). The average Beta and average price inverse are slightly lower in this paper than in Hong & Kacperczyk (2009). See appendix table 11 for the correlation matrix of all variables.

The variables Rightdum and Sindum2 utilize data on MSCI ESG factors. The availability of data of these factors is more scarce compared to the other variables. I create a sub dataset from the previous dataset of 81,662 obeservations that do not have missing values for Rightdum and Sindum2. This dataset consists of 24,983 observations and 3,870 different firms. The summary statistics of this sub dataset are also included in table 2 and are displayed on every second row with 'N= 24,983'. The sub dataset includes relatively more large S&P 500 firms with more institutional ownership. The mean institutional ownership is 0.694 and the mean Logsize and SP500 are 21.206 and 0.255 respectively.

Table 2
Summary statistics of dependent and control variables

This table contains summary statistics of the dependent variable institutional ownership and the control variables for years 1990 to 2016. The total sample consists of 81,662 observations. The sub dataset using ESG factors consists of 24,983 observations. Percentage institutional ownership is defined as shares owned of stock i by institutional investors for year t divided by the total number of shares outstanding of stock i for year t. LOGSIZE is the measure for size and is calculated by taking the natural logarithm of market capitalization. Market capitalization is calculated as end of year price on year t for firm i times shares outstanding on end of year t for firm i. LOGMB is the natural logarithm of the market to book value of firm i for year t and is measured as the end of year market capitalization divided by the end of year book value. The variable SP500 is a dummy variable that takes on the value one if the stock is listed on the S&P 500 index and zero otherwise. The rather similar variable NASD is a dummy variable that takes on the value one if the stock is listed on the NASDAQ and zero otherwise. STD is the standard deviation of daily returns for firm i during year t. BETA is the year-end beta reported by Compustat in year t for firm i. PRINV is the inverse of the firm i share price at the end of year t. RET is the average monthly return on stock i during year t.

	N	Minimum	Maximum	Mean	St. Dev
Inst. own. (%)	81,662	0.000	1	0.430	0.297
	24,983	0.000	1	0.694	0.207
Logsize ('000)	81,662	10.427	29.834	19.631	2.270
	24,983	14.954	27.188	21.206	1.613
LogMB	81,662	-8.835	10.711	0.771	0.962
	24,983	-3.621	8.638	0.952	0.808
STD	81,662	0	1.208	0.036	0.024
	24,983	0.004	0.397	0.026	0.013
RET	81,662	-0.409	1.176	0.013	0.052
	24,983	-0.227	1.176	0.013	0.038
Beta	81,662	-3.753	5.934	0.879	0.654
	24,983	-0.901	5.830	1.148	0.560
Prinv	81,662	0.000	52.632	0.205	0.726
	24,983	0.000	6.289	0.066	0.136
NASD	81,662	0	1	0.438	0.496
	24,983	0	1	0.428	0.495
SP500	81,662	0	1	0.101	0.301
	24,983	0	1	0.255	0.435

Graph 1
Average institutional ownership per year for years 1990 to 2016

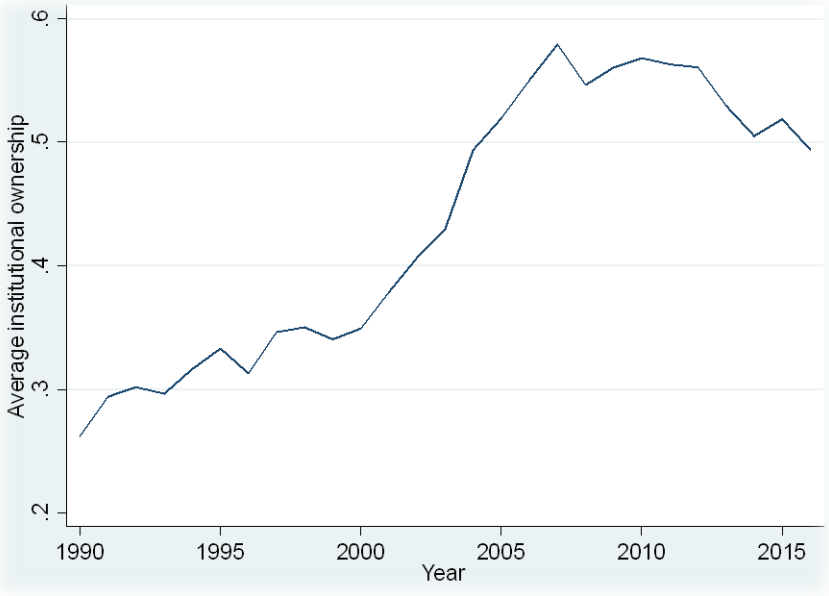


Table 3 shows the mean and standard deviation of the different dummy variables. The mean can be interpreted as the proportion of '1's' in the sample. Take for example Sindum1; 1.2% of the entire sample is a sin stock[11]. The tobacco sin dummy contains the lowest amount of sin stocks, namely 0.1 %. The alcohol and gambling dummies are better represented and have a 0.5% and 0.6% representation respectively. Using the second sin stock selection method 8.7% of the sample is a sin stock and 10.1% of the sample is a righteous stock. The sample size for the last two is smaller due to availability of MSCI KLD data which leads to a higher mean for these two variables. Table 3 also contains an overview of the number of sin stocks (per industry) and righteous stocks per year. The number of sin stocks in the variable Sindum1 ranges between 16 and 52 and the total amount of observations is 968. The number of sin stocks for the other Sin dummy (Sindum2) is generally larger, especially for later years. The year 2016 however only contains 42 sin stocks. This is odd given the pattern of Sindum2 over the years. Diving into the data and database manuals provides no clear explanation for this particular finding. The Sindum2 variable ranges between 16 and 169 and has a total of 2196 sin stock observations. Alcdum, Tobdum and Gamdum are relatively stable over the years. Their total number of observations are 87, 370 and 512 respectively. The righteous stock dummy has considerably more righteous stock in later years than in earlier years. The numbers vary between 40 and 306 with a total of 2739 observations. See appendix table 12 for a similar overview of stocks per year for the factors making up the second sin dummy and the righteous stock dummy.

Table 3

Summary statistics of the indepent variables of interest

This table includes descriptive statistics of the indepent variables of interest. The total number of observations is 81,662 for the first four columns and 24.983 for the last two. This table presents the number of firms denoted as Sindum1, Alcdum, Tobdum, Gamdum, Sindum2 and Rightdum by year for years 1990 to 2016. The last two rows report the mean and the standard deviation of the mean. The mean can be interpreted as the proportion of stocks denoted, as for example Sindum1, as a proportion of the total sample. The data for Sindum2 and Rightdum is available from 1995 to 2016, hence the years 1990 to 1994 are missing. The minimum and maximum of all variables are respectively 0 and 1. Sindum1 is dummy variable that takes on the value one if it falls in the alcohol tobacco or gambling industry and is zero otherwise. Alcdum is a dummy variable that is equal to one if the stock is an alcohol sin stock (SIC codes 2080-2085) and zero otherwise. Gamdum is a dummy variable that is equal to one if the stock is a gambling sin stock (NAICS codes 7132, 71312, 713210, 71329, 713290, 72112, and 721120) and zero otherwise. Tobdum is a dummy variable that is equal to one if the stock is a tobacco sin stock (SIC codes 2100-2199) and zero otherwise. Sindum2 is a dummy variable that takes on the value 1 if it is denoted as 1 on any of the MSCI KLD ESG factors human rights, child labor, waste management and climate change. Rightdum is a dummy variable that takes on the value 1 if it is denoted as 1 on any of the positive counterpart MSCI KLD ESG factors human rights, waste management and climate change. Total n stands for the total number of observations per column. Total N is total number of observations of the entire sample.

[11] As illustration: 968/81,662=0.012. 968 is taken from total N of table 3 and 81,662 is the total number of observations.

Year	Sindum1	Tobdum	Alcdum	Gamdum	Sindum2	Rightdum
1990	17	2	6	9	n/a	n/a
1991	16	2	5	9	n/a	n/a
1992	19	2	6	11	n/a	n/a
1993	29	2	7	20	n/a	n/a
1994	34	2	7	25	n/a	n/a
1995	40	3	8	29	17	40
1996	51	3	15	33	26	54
1997	52	5	16	31	27	58
1998	46	4	17	25	32	61
1999	43	3	17	23	51	57
2000	41	2	18	21	50	49
2001	35	2	15	18	92	57
2002	39	2	17	20	77	60
2003	42	3	18	21	106	77
2004	44	4	18	22	147	54
2005	40	5	15	20	126	53
2006	35	4	14	18	152	65
2007	34	3	13	18	159	84
2008	35	5	12	18	168	96
2009	34	5	12	17	168	97
2010	34	4	14	16	126	270
2011	34	4	15	15	118	250
2012	33	3	16	14	105	171
2013	35	3	18	14	159	238
2014	34	3	17	14	122	184
2015	35	3	17	15	106	222
2016	37	4	17	16	42	223
Total n	968	87	370	512	2176	2520
Total N	81,662	81,662	81,662	81,662	24,983	24,983
Mean	0.012	0.001	0.005	0.006	0.087	0.101
Mean STD	0.108	0.033	0.067	0.079	0.279	0.301

4. Methodology

The methodology section is subdivided into four parts that correspond to the four hypotheses. The first part deals with the general relationship between sin and righteous stocks and institutional ownership. The second section focusses on the differences of the effect of the sin stock dummy across time periods. The third part focusses on the differences between the three different sin stock industries and the fourth and final part focusses on the differences between different types of institutional investors.

The first section forms the base of this paper and focusses on the relationship between institutional ownership, righteous stocks and sin stocks. This part is designed to answer hypotheses 1A and 1B. The

first linear multivariate regression is executed as a whole on all panel data. Hypotheses 1A and 1B are answered by running the regression as a whole because this gives the general difference between the proportion of institutional ownership of (non-)sin stocks and (non-)righteous stocks. This regression allows me to answer whether institutions hold a larger proportion of shares of non-sin stock in comparison to sin stocks. Additionally this section allows me to assess whether righteous stock have a larger proportion of institutional ownership in comparison to non-righteous stock. The variable Sindum is constructed twice (explained in section 3.2), hence the regression is estimated twice using the two separate methods of selecting sin stocks (traditional industry way and MSCI ESG factors way). The OLS regression contains the dependent variable Percentage institutional ownership and independent variables Sindum, Rightdum and a vector of firm characteristics.

The first linear multivariate regression is as follows:

$$IO_{it} = \alpha + \beta_1 SINDUM_{it} + \beta_2 RIGHTDUM_{it} + \beta_3 X_{it} + \varepsilon_{it} \quad i = 1, \dots, N$$

Sindum is a dummy variable that is equal to one if the stock is a sin stock and zero otherwise. Rightdum is a dummy variable that is equal to one if the stock is a righteous stock and zero otherwise. X_{it} is a vector of firm characteristics including Onedigdum/ Ffdum, size, beta, market to book value, standard deviation of daily returns, inverse of firm i's share price, average monthly return, enlistment on NASD, enlistment on S&P 500 and ε_{it} is the error term.

The statistical hypotheses for regression one are as follows:

H0: $\beta_1 = 0, \beta_2 = 0,$	The coefficients β_1, β_2 are not significantly different from zero.
H1: $\beta_1 \neq 0, \beta_2 \neq 0$	At least one of the coefficients β_1, β_2 is significantly different from zero.

The second methodology section focusses on hypothesis 2. This part utilizes a similar regression model as in part 1. However, for this section I run pooled regressions, whereby the sample is subdivided into different time periods. The sample is subdivided into time periods of five years and one time period of seven years (the last time period). Running the regression across different time periods provides the difference of the effect of the sin stock dummy on percentage institutional ownership between time periods. This is necessary to answer the question of whether the difference in proportion of institutional ownership between sin stocks and non-sin stocks has become larger over time (hypothesis 2). This section is executed using the traditional sin stock selection method and without Rightdum. These regressions will provide a Sindum coefficient for every time period. Whether these coefficients are significantly different is tested using a chi-square test. In short, these regressions deliver the behavior of the sin stock dummy over time.

The third methodology section contains a linear multivariate regression for the different types of sin stocks. This section is designed with the purpose of finding out whether the proportion of institutional ownership of tobacco, gambling and alcohol firms differ. This part is designed to answer hypothesis 3. The coefficients β_1, β_2 and β_3 are of interest in the regression shown below. These coefficients show whether there is a difference in the proportion of institutional ownership for the different sin stock industries.

The third linear multivariate regression is comparable to the first regression and is as follows:

$$IO_{it} = \alpha + \beta_1 TOBDUM_{it} + \beta_2\, GAMDUM_{it} + \beta_3\, ALCDUM_{it} + \beta_4 X_{it} + \varepsilon_{it} \quad i = 1, \dots, N$$

TOBDUM is a dummy variable that is equal to one if the stock is a tobacco sin stock and zero otherwise. GAMDUM is a dummy variable that is equal to one if the stock is a gambling sin stock and zero otherwise. ALCDUM is a dummy variable that is equal to one if the stock is an alcohol sin stock and zero otherwise. X_{it} is a vector of firm characteristics including Onedigdum/ Ffdum, size, beta, market to book value, standard deviation of daily returns, inverse of firm i's share price, average monthly return, enlistment on NASD and enlistment on S&P 500 and ε_{it} is the error term.

The statistical hypotheses for regression three are as follows:

H0: $\beta_1 = 0, \beta_2 = 0, \beta_3 = 0$ The coefficients $\beta_1, \beta_2, \beta_3$ are not significantly different from zero.

H1: $\beta_1 \neq 0, \beta_2 \neq 0, \beta_3 \neq 0$ At least one of the coefficients $\beta_1, \beta_2, \beta_3$ is significantly different from zero.

The size, sign and significance of coefficients $\beta_1, \beta_2\ and\ \beta_3$ will be known after this regression. However, it is not clear yet whether coefficients $\beta_1, \beta_2\ and\ \beta_3$ are significantly different from each other. Whether these coefficients are significantly different is tested using an F-test.

The fourth and final methodology section focusses on differences between different types of institutional investors and answers hypothesis 4 A and B. In order to test whether there are differences between more and less norm-constrained investors I create two groups of institutional investors. The first group includes shareholder types 1, 2 and 5 (bank, insurance company and all others including pension plans, university endowment funds and religious organizations). Shareholder type 5 is excluded from the data starting from the last quarter of 1997 due to the measurement error described in section 3.3[12]. The second group includes shareholder types 3 and 4 (investment companies and their managers and independent investment advisor). Next, I construct two new dependent variables based

[12] Hypothesis 4 is also estimated without the exclusion of shareholder type 5. The results are discussed in the results section of hypothesis 4.

on the two groups. Similarly as for hypothesis 1, I estimate this regression twice; once using the traditional sin stock selection method and once using both positive and negative MSCI KLD ESG factors. The following linear multivariate regressions are estimated:

$$IOGROUP1_{it} = \alpha + \beta_1 SINDUM_{it} + \beta_2 RIGHTDUM_{it} + \beta_3 X_{it} + \varepsilon_{it} \;\; i = 1, \dots, N$$

$$IOGROUP2_{it} = \alpha + \beta_1 SINDUM_{it} + \beta_2 RIGHTDUM_{it} + \beta_3 X_{it} + \varepsilon_{it} \;\; i = 1, \dots, N$$

The description of the variables is similar to the description under the first methodology part (hypothesis 1). The only difference is the dependent variable. The dependent variables stand for the percentage of institutional ownership for group 1 (regression 1) and the percentage of institutional ownership for group 2 (regression 2). The two important coefficients in these regressions are β_1 and β_2. These coefficients will show whether there is a difference in sin stock ownership between more and less norm-constrained institutional investors. The significance of this difference is tested using a chi-square test.

The statistical hypotheses for this methodology section are as follows:

H0: $\beta_1(IOGROUP1) = \beta_1(IOGROUP2),\ \beta_2(IOGROUP1) = \beta_2(IOGROUP2)$

The coefficients β_1, β_2 of group 1 and 2 are not significantly different from each other.

H1: $\beta_1(IOGROUP1) \neq \beta_1(IOGROUP2),\ \beta_2(IOGROUP1) \neq \beta_2(IOGROUP2)$

At least one of the coefficients β_1, β_2 of group 1 and 2 is significantly different from each other.

All multivariate regression models use panel data and therefore it is important to be aware of covariance between unobserved heterogeneity and one or more of the explanatory variables. This covariance has to be zero in order for OLS to be consistent. This problem will be, at least partly, tackled using fixed effects. Additional regressions are estimated with year fixed effects to show robustness of the results. The idea behind this is to control for individual effects that are unique to a particular year (graph 1 has shown a changing average institutional ownership over time), thereby taking out the time trend that is not captured in the control variables. The standard errors are another important aspect of estimating results in a statistically correct manner. The standard errors in the regressions are clustered at the firm-level[13]. This addresses the potential issue of errors being correlated within firms over time. This is a more conservative measure than using for example White standard errors. Clustering at the firm-level means that observations from a single firm are not seen as independent

[13] Another method for clustering standard errors is through clustering at the industry-level. The results are estimated using industry clustering as well and the results are discussed in the robustness check section.

and unrelated observations. This decreases the statistical information that can be drawn from these observations. This measure is therefore more conservative and will be used throughout this paper.

5. Results

The results section follows the order of the hypotheses and methodology. The first section lays out the results on the first hypothesis. The results in table 4 show that there is a negative relationship between sin stocks and institutional ownership, at least for the traditional way of selecting sin stock. The initial effect of the sin stock dummy on institutional ownership is -0.0805 and is significant at the 1% level. This effect becomes more negative when control variables are added. The coefficient of the sin stock dummy is -0.1183 for regression 8 which means that sin stocks have an 11.83 percentage points lower institutional ownership in comparison to non-sin stocks. The relationship is highly significant at the 1% level. The relationship is not only statistically significant but also economically significant. The average institutional ownership is 43% so 11.54% is more than a quarter of the average. The relationship stays negative and significant across different combinations of control variables and when including year fixed effects. In column 2, Onedigdum has a value of 0.0015, however this value is not significant at the 10% level. In column 3, the value for Ffdum is 0.0108 and is not significant at the 10% level either. Both variables control for comparable industries so for further regressions only one of the two is included. What matters for these variables is how it changes the value of the Sindum coefficient when added to the regression. Onedigdum appears to make the Sindum variable more positive so I therefore stick to Onedigdum for further regressions. The effect of this is that the effect of the sin stock dummy on institutional ownership becomes smaller, which works against the hypothesis. This reasoning also stands for further regressions. All other control variables, but Onedigdum and Prinv are highly significant at the 1 % level for regression 8. Logsize, Beta and return have a positive effect on institutional ownership which means that firms that are larger and firms that have a higher beta have a higher proportion of institutional ownership. LogMB, Prinv and Std have a negative relationship with percentage institutional ownership. Lastly, NASD and S&P 500 have a positive effect on the percentage of institutional ownership. Adding year fixed effects, as shown in column 9, does not change these results in general. These results are in accordance with hypothesis one which stated that the proportion of institutional ownership of sin stocks is lower than the proportion of institutional ownership of non-sin stocks.

The initial value of the second sin dummy is -0.0067 and is not significant at the 10% level (see table 5). This value becomes more negative as more control variables are added to the regression until it varies between roughly -0.03 and -0.04 for the more complete regressions 7 and 8. The variables

Onedigdum and Ffdum are not added to this regression because they are not informative. This sin stock selection method selects stocks across different industries and not just from the three traditional sin industries (alcohol, tobacco and gambling). These two variables are therefore not informative. The findings in table 5 are in line with the findings in table 4, regarding sin stocks. The variable Sindum2 in table 5 has a negative coefficient, similarly as in table 4. This means that sin stocks have a lower proportion of institutional ownership than non-sin stocks. This result is also in accordance with hypothesis one which stated that the proportion of institutional ownership of sin stocks is lower than the proportion of institutional ownership of non-sin stocks. The economic significance of Sindum2 is not as large as for Sindum1. It is important to be aware of the fact that the sin stock selection process is different from the sin stock selection process for sindum1 and that the sample size of the regressions differ. The difference in sample size along with difference in selection factors potentially explain the difference in magnitude of the coefficients. Further explanations of the results will be discussed in the discussion section.

The initial value of the righteous stock dummy is -0.0451 and is significant at the 1% level. The coefficient grows slightly more negative as control variables are added and ends up with a coefficient of approximately -0.07. The relationship between the righteous stock dummy and institutional ownership remains highly significant at the 1% level across different combinations of variables. Beside statistical significance, the coefficient is also economically significant when compared to the mean of institutional ownership (0.43). What can be taken from these regressions is that institutional investors do not seem to hold a larger proportion of righteous stock than of non-righteous stock. Hypothesis 1B predicted that the proportion of institutional ownership of righteous stock is higher than the proportion of institutional ownership of non-righteous stock. This is not the case for the regressions because of the negative coefficient for Rightdum. The results are therefore not in line with the hypothesis.

Table 4
Institutional ownership and sin stocks

This table contains the first regression including the dependent variable institutional ownership and the control variables for years 1990 to 2016. The total sample consists of 81,662 observations. Percentage institutional ownership is defined as shares owned of stock i by institutional investors for year t divided by the total number of shares outstanding of stock i for year t. Sindum1 is dummy variable that takes on the value one if it falls in the alcohol tobacco or gambling industry and is zero otherwise. ONEDIGDUM is a dummy variable which takes on the value one if the one digit sic code coincides with the one digit sic code of the alcohol, tobacco or gambling industry (codes 2 and 7), and is zero otherwise. FFDUM is a dummy variable that takes on the value one if the stock falls into one of the following Fama and French (1997) 49 industry categories 2 (food), 3 (soda), 4 (beer), 5 (smoke), 7 (fun), 43 (meals), and 49 (gambling), and zero otherwise. LOGSIZE is the measure for size and is calculated by taking the natural logarithm of market capitalization. Market capitalization is calculated as end of year price on year t for firm i times shares outstanding on end of year t for firm i. LOGMB is

the natural logarithm of the market to book value of firm i for year t and is measured as the end of year market capitalization divided by the end of year book value. The variable SP500 is a dummy variable that takes on the value one if the stock is listed on the S&P 500 index and zero otherwise. The rather similar variable NASD is a dummy variable that takes on the value one if the stock is listed on the NASDAQ and zero otherwise. STD is the standard deviation of daily returns for firm i during year t. BETA is the year-end beta reported by Compustat in year t for firm i. PRINV is the inverse of the firm i share price at the end of year t. RET is the average monthly return on stock i during year t. The numbers in brackets are the standard errors that are robust to heteroscedasticity and are clustered within firm over time. *** 1% level of significance; ** 5% level of significance; and * 10% level of significance.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Sin-dum1	-.0805***	-.0815***	-.0901***	-.1256***	-.1136***	-.1149***	-.1194***	-.1183***	-.1172***
	(.0284)	(.0288)	(.0300)	(.0302)	(.0289)	(.0277)	(.0269)	(.0250)	(.0247)
Onedigdum		.0015		-.0045	-.0036	.0072	.0111*	.0048	.0032
		(.0071)		(.0063)	(.0061)	(.0062)	(.0060)	(.0059)	(.0058)
Ffdum			.0108						
			(.0109)						
Logsize ('000)				.0621***	.0550***	.0585***	.0463***	.0433***	.0314***
				(.0018)	(.0018)	(.0020)	(.0024)	(.0029)	(.0032)
Beta					.0772***	.0794***	.0940***	.0967***	.0902***
					(.0031)	(.0031)	(.0034)	(.0034)	(.0037)
LogMB						-.0399 ***	-.0358***	-.0385***	-.0304***
						(.0028)	(.0029)	(.0028)	(.0029)
Prinv						-.0305***	-.0100**	-.0050	-.0091***
						(.0047)	(.0044)	(.0043)	(.0051)
STD							-2.198***	-2.339***	-2.618***
							(.1554)	(.1639)	(.2072)
RET							.1580***	.1572***	.2820***
							(.0216)	(.0215)	(.0240)
NASD								.0867***	.0712***
								(.0052)	(.0052)
SP500								.0976***	.1227***
								(.0118)	(.0117)
Year FE									yes
R2	0.0009	0.0009	0.0010	0.2252	0.2511	0.2689	0.2876	0.3126	0.3472
#obs	81,662	81,662	81,662	81,662	81,662	81,662	81,662	81,662	81,662

Table 5

Institutional ownership, righteous stock and sin stock

This table contains the first regression including the dependent variable institutional ownership and the control variables for years 1990 to 2016. The total sample consists of 24,983 observations. Percentage institutional ownership is defined as shares owned of stock i by institutional investors for year t divided by the total number of shares outstanding of stock i for year t. Sindum2 is a dummy variable that takes on the value 1 if it is denoted as 1 on any of the MSCI KLD ESG factors human rights, child labor, waste management and climate change. Rightdum is a dummy variable that takes on the value 1 if it is denoted as 1 on any of the positive counterpart MSCI KLD ESG factors human rights, child labor, waste management and climate change. LOGSIZE is the measure for size and is calculated by taking the natural logarithm of market capitalization. Market capitalization is calculated as end of year price on year t for firm i times shares outstanding on end of year t for firm i. LOGMB is the natural logarithm of the market to book value of firm i for year t and is measured as the end of year market capitalization divided by the end of year book value. The variable SP500 is a dummy variable that takes on the value one if the stock is listed on the S&P 500 index and zero otherwise. The rather similar variable NASD is a dummy variable that takes on the value one if the stock is listed on the NASDAQ and zero otherwise. STD is the standard deviation of daily returns for firm i during year t. BETA is the year-end beta reported by Compustat in year t for firm i. PRINV is the inverse of the firm i share price at the end of year t. RET is the average monthly return on stock i during year t. The numbers in brackets are the standard errors that are robust to heteroscedasticity and are clustered within firm over time. *** 1% level of significance; ** 5% level of significance; and * 10% level of significance.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sindum 2	-.0067	-.0261**	-.0313***	-.0368***	-.0337***	-.0330***	-.0279***	-.0433***
	(.0099)	(.0103)	(.0100)	(.0101)	(.0099)	(.0396)	(.0101)	(.0101)
Right-dum	-.0451***	-.0808***	-.0751***	-.0790***	-.0696***	-.0704***	-.0660***	-.0696***
	(.0076)	(.0079)	(.0077)	(.0078)	(.0077)	(.0076)	(.0075)	(.0074)
Logsize ('000)		.0196***	.0208***	.0235***	.0147***	.0118***	.0182***	.0154***
		(.0024)	(.0023)	(.0025)	(.0027)	(.0027)	(.0036)	(.0038)
Beta			.0424***	.0424***	.0479***	.0581***	.0589***	.0604***
			(.0047)	(.0047)	(.0046)	(.0047)	(.0047)	(.0051)
LogMB				-.0132***	-.0174***	-.0136***	-.0167***	-.0112***
				(.0037)	(.0036)	(.0037)	(.0038)	(.0038)
Prinv					-.2845***	-.2713***	-.2614***	-.2524***
					(.0396)	(.0405)	(.0399)	(.0400)
STD						-.9583***	-1.033***	-2.467***
						(.2197)	(.2239)	(.3327)
RET						-.2320***	-.2459***	.0744
						(.0423)	(.0431)	(.0537)
NASD							.0216***	.0137*
							(.0071)	(.0071)
SP500							-.0231**	.0045
							(.0094)	(.0097)

Year FE								yes
R2	0.0046	0.0240	0.0370	0.0393	0.0688	0.0732	0.0768	0.1569
#obs	24,983	24,983	24,983	24,983	24,983	24,983	24,983	24,983

Table 6 presents the results of the regression analysis that is designed to answer hypothesis two. These regressions are constructed to assess the behaviour of the sin stock dummy over time. The data time span of 27 years is subdivided into five groups. The five groups are: (1) 1990-1994, (2) 1995-1999, (3) 2000-2004, (4) 2005- 2009 and (5) 2010-2016. In table 6, the sindum1 coefficient becomes more negative over time. The coefficent is -0.0198 for the first time period and -0.1818 for the last time period, whereby the first time period is not statistically different from zero and the last time period is highly significant at the 1% level. Onedigdum is used for these regressions instead of Ffdum for similar reasons as in regression one. Panel B reports whether the coefficients of Sindum1 are significantly different from each other. All tests except for test 2 report a positive difference which means that the sin stock dummy has become more negative in the latter time period. The differences between time periods 1 and 2 and 1 and 5 are significant at respectively the 5% and 1% level. This means that these coefficients are significantly different from each other. Tests 2, 3 and 4 in panel B are not significant however. The results show that the proportion of institutional ownership of sin stocks significantly changes when comparing time periods far apart. The more adjacent time periods do not have a significantly different coefficient for Sindum, except for the first two time periods. The following conclusion follows after following the general pattern in the sin stock dummy and judging from the highly significant difference between time period 1 and 5. The effect of the sin stock dummy has become signifcantly larger over time. Appendix table 13 reports all sin stock dummies for all years (1990-2016). The sindum1 coefficient is not significant for the first 3 years and is highly significant at the 1 % level for the last 10 years. In the beginning the coefficient is not significantly different from zero and followed by years whereby the coefficient varies between -0.05 and -0.10. After 2007 the Sindum effect seems even larger with coefficients between -0.12 and -0.22. Therefore this appendix table also shows that the effect of the sin stock dummy has become larger over time. One example of a sin stock that reports this behavior is the Altria Group INC stock, which is a tobacco sin stock. From 2003 to 2016 the percentage institutional ownership of this stock followed the following pattern: 0.6207, 0.7143, 0.7435, 0.7170, 0.7007, 0.6507, 0.6554, 0.5899, 0.5688, 0.5575, 0.5011, 0.4526, 0.4696 and 0. 4796. Another example from the alcohol industry which shows a similar pattern is the Brown Forman CORP. The Brown Forman CORP stock starts off with institutional ownership of 70% in 1998 and ends with institutional ownership of 40% in 2016. The findings combined find evidence in

favor of hypothesis two that stated that the difference in proportion of institutional ownership between sin stocks and non-sin stocks becomes larger over time.

Table 6

Sin stock institutional ownership over time

This table contains an overview of the regression coefficients of the variable Sindum for five different time periods between the year 1990 and 2016. The five time periods are: (1) 1990-1994, (2) 1995-1999, (3) 2000-2004, (4) 2005-2009 and (5) 2010-2016. The total sample consists of 81,662 observations. The dependent variable percentage institutional ownership is defined as shares owned of stock i by institutional investors for year t divided by the total number of shares outstanding of stock i for year t. Sindum1 is dummy variable that takes on the value one if it falls in the alcohol tobacco or gambling industry and is zero otherwise. ONEDIGDUM is a dummy variable which takes on the value one if the one digit sic code coincides with the one digit sic code of the alcohol, tobacco or gambling industry (codes 2 and 7), and is zero otherwise. LOGSIZE is the measure for size and is calculated by taking the natural logarithm of market capitalization. Market capitalization is calculated as end of year price on year t for firm i times shares outstanding on end of year t for firm i. LOGMB is the natural logarithm of the market to book value of firm i for year t and is measured as the end of year market capitalization divided by the end of year book value. The variable SP500 is a dummy variable that takes on the value one if the stock is listed on the S&P 500 index and zero otherwise. The rather similar variable NASD is a dummy variable that takes on the value one if the stock is listed on the NASDAQ and zero otherwise. STD is the standard deviation of daily returns for firm i during year t. BETA is the year-end beta reported by Compustat in year t for firm i. PRINV is the inverse of the firm i share price at the end of year t. RET is the average monthly return on stock i during year t. The numbers in brackets are the standard errors that are robust to heteroscedasticity and are clustered within firm over time. *** 1% level of significance; ** 5% level of significance; and * 10% level of significance.

Panel A: regression analysis					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Sindum1	-.0198	-.0919***	-.0854**	-.1301***	-.1818***
	(.0323)	(.0228)	(.0358)	(.0376)	(.0377)
Onedigdum	-.0046	.0037	.0007	.0238**	-.0125
	(.0064)	(.0062)	(.0081)	(.0098)	(.0086)
Logsize ('000)	.0417***	.0433***	.0267***	.0285***	.0278***
	(.0032)	(.0033)	(.0035)	(.0047)	(.0042)
Beta	.0690***	.0653***	.1192***	.1073***	.0988***
	(.0038)	(.0041)	(.0057)	(.0077)	(.0064)
LogMB	-.0337***	-.0475***	-.0416***	-.0730***	.0212***
	(.0033)	(.0034)	(.0043)	(.0053)	(.0048)
Prinv	.0057***	.0000	-.027***	-.0624***	-.0214**
	(.0024)	(.0120)	(.0092)	(.0157)	(.0094)
STD	-1.307***	-2.122***	-2.914***	-2.465***	-3.663***
	(.1456)	(.1715)	(.1540)	(.6120)	(.4224)
RET	.2191***	.3804***	.2129***	.0424	.2319***
	(.0286)	(.0330)	(.0342)	(.0618)	(.0594)

NASD	.0318***	.0422***	.0907***	.1102***	.0687***
	(.0056)	(.0056)	(.0077)	(.0090)	(.0080)
SP500	.0991***	.1112***	.1471***	.1651***	.0771***
	(.0134)	(.0137)	(.0147)	(.0162)	(.0147)
R2	0.3793	0.3159	0.2908	0.2344	0.2444
#obs	13,138	18,019	16,085	14,730	19,690

Panel B: chi-square test analysis of difference of coefficients

Combinations of variables	**Difference**
Sindum1 (1) - Sindum1 (2)	0.0721**
Sindum1 (2) - Sindum1 (3)	-0.0065
Sindum1 (3) - Sindum1 (4)	0.0447
Sindum1 (4) -Sindum1 (5)	0.0517
Sindum1 (1) - Sindum1 (5)	0.1620***

The next results section contains the analysis of hypothesis 3 whereby the difference in institutional ownership for different sin industries is tested. Panel A in table 7 presents the regression analysis. Column 1 shows that all sin industry dummies have a negative relationship with institutional ownership. The alcohol and gambling dummy have a 5% significance level, but the tobacco dummy is not significant. The significance of the alcohol and gambling dummy increase to the 1% level after adding the control variables. The coefficient for the tobacco industry remains not significant at the 1% level. The size of the coefficients are roughly -0.15 for the tobacco and alcohol industry for regression 7. The gambling industry has a different size which is roughly -0.08. The interpretation for these coefficients is similar as to previous regressions. For instance alcohol sin stock have a 15 percentage points lower institutional ownership in comparison to non-alcohol industry stocks. The coefficients of the sin industry dummies are also economically significant judging from the values in regression 7 and comparing their magnitude to the mean institutional ownership (0.43). Adding year fixed effects does not make a significant change to these findings. What remains now is testing whether the coefficients of the different sin industries are significantly different from each other. This is tested using an F-test. The results of this F-test are displayed in Panel B of table 7. None of the F-tests show a significant difference. This means that the relationship that these industries have with institutional ownership are not significantly different from each other. The last test tests the null hypothesis that all coefficients are the same. The joint test of all three coefficients being the same cannot be rejected. Hypothesis 3 predicted that the different sin industries have a significantly different relationship with the proportion of institutional ownership. This does not hold for any of the combinations of sin industries. The results are therefore not in line with the hypothesis.

Table 7

Institutional ownership and the tobacco, alcohol and gambling industry

This table contains the results of a regression analysis and an F-test. The regression includes the dependent variable institutional ownership and control variables for years 1990 to 2016. The total sample consists of 81,662 observations. Percentage institutional ownership is defined as shares owned of stock i by institutional investors for year t divided by the total number of shares outstanding of stock i for year t. Alcdum is a dummy variable that is equal to one if the stock is an alcohol sin stock (SIC codes 2080-2085) and zero otherwise. Gamdum is a dummy variable that is equal to one if the stock is a gambling sin stock (NAICS codes 7132, 71312, 713210, 71329, 713290, 72112, and 721120) and zero otherwise. Tobdum is a dummy variable that is equal to one if the stock is a tobacco sin stock (SIC codes 2100-2199) and zero otherwise. ONEDIGDUM is a dummy variable which takes on the value one if the one digit sic code coincides with the one digit sic code of the alcohol, tobacco or gambling industry (codes 2 and 7), and is zero otherwise. LOGSIZE is the measure for size and is calculated by taking the natural logarithm of market capitalization. Market capitalization is calculated as end of year price on year t for firm i times shares outstanding on end of year t for firm i. LOGMB is the natural logarithm of the market to book value of firm i for year t and is measured as the end of year market capitalization divided by the end of year book value. The variable SP500 is a dummy variable that takes on the value one if the stock is listed on the S&P 500 index and zero otherwise. The rather similar variable NASD is a dummy variable that takes on the value one if the stock is listed on the NASDAQ and zero otherwise. STD is the standard deviation of daily returns for firm i during year t. BETA is the year-end beta reported by Compustat in year t for firm i. PRINV is the inverse of the firm i share price at the end of year t. RET is the average monthly return on stock i during year t. The numbers in brackets are the standard errors that are robust to heteroscedasticity and are clustered within firm over time. *** 1% level of significance; ** 5% level of significance; and * 10% level of significance.

Panel A: Regression analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alcdum	-.1148**	-.1159**	-.1806***	-.1516***	-.1494***	-.1537***	-.1562***	-.1622***
	(.0585)	(.0587)	(.0593)	(.0575)	(.0541)	(.0529)	(.0473)	(.0460)
Gamdum	-.0667**	-.0677**	-.0675**	-.0736***	-.0826***	-.0884***	-.0851***	-.0781***
	(.0287)	(.0290)	(.0265)	(.0267)	(.0270)	(.0258)	(.0259)	(.0259)
Tobdum	-.0153	-.0163	-.2352**	-.1890*	-.1587	-.1572	-.1530	-.1573
	(.0942)	(.0943)	(.1138)	(.1096)	(.1120)	(.1082)	(.0966)	(.0970)
Onedig-dum		.0015	-.0045	-.0035	.0071	.0111*	.0047	.0031
		(.0071)	(.0063)	(.0061)	(.0062)	(.0060)	(.0059)	(.0058)
Logsize ('000)			.0622***	.0551***	.0586***	.0463***	.0434***	.0315***
			(.0018)	(.0018)	(.0020)	(.0024)	(.0029)	(.0032)
Beta				.0769***	.0792***	.0937***	.0965***	.0899***
				(.0031)	(.0031)	(.0034)	(.0034)	(.0037)
LogMB					-.0398***	-.0358***	-.0385***	-.0303***
					(.0028)	(.0029)	(.0028)	(.0029)
Prinv					-.0304***	-.0100**	-.0050	-.0091*
					(.0047)	(.0044)	(.0043)	(.0051)

STD						-2.199***	-2.340***	-2.619***
						(.1554)	(.1639)	(.2071)
RET						.1574***	.1568***	.2816***
						(.0216)	(.0214)	(.0240)
NASD							.0865***	.0709***
							(.0052)	(.0052)
SP500							.0980***	.1233***
							(.0118)	(.0117)
Year FE								yes
R2	0.0010	0.0010	0.2257	0.2514	0.2691	0.2878	0.3128	0.3475
#obs	81,662	81,662	81,662	81,662	81,662	81,662	81,662	81,662

Panel B: F-test analysis of difference of coefficients

Combinations of variables	**Difference (column 7)**
Alcdum - Tobdum	-0.0031
Alcdum - Gamdum	-0.0711
Tobdum - Gamdum	-0.0679
Joint F–test	**P-value**
Gamdum = Alcdum = Tobdum	0.3622

The results on the final hypothesis are presented in table 8. The coefficient for more norm-constrained institutional investors (shareholder types 1,2 and 5) is -0.0149 and the coefficient for less norm-constrained institutional investors is -0.0355. The former is significant at the 5% level and the latter is significant at the 1% level. Columns 3 and 4 with year fixed effects include similar results. The findings mean that less norm-constrained investors hold a relatively lower proportion of sin stocks than more norm-constrained investors. Panel B reports the results of the test whereby the significance of the difference of the coefficients of Sindum1 between both regressions is tested. The difference is 0.0206 and the coefficients are significantly different at the 5% level. Hypothesis 4 stated that the proportion of sin stock ownership of more norm-constrained institutional investors is relatively lower than the proportion of sin stock ownership of less norm-constrained institutional investors. The results in table 8 reject this because the coefficient for more norm-constrained investors is significantly less negative than for less norm-constrained investors. These results therefore provide evidence opposing the social norm hypothesis.

Table 8

Sin stock ownership of more and less norm-constrained institutional investors

This table presents a pooled regression including two dependent variables and control variables for years 1990 to 2016. The total sample consists of 80,851 observations. Proportion institutional ownership for more norm-constrained investors is calculated by summing up the total shares owned by shareholder types 1, 2 and 5 for year t and dividing by total number of shares outstanding for year t. Proportion institutional ownership for less norm-constrained investors is calculated by summing up the total shares owned by shareholder types 3 and 4 for year t and dividing by total number of shares outstanding for year t. Sindum1 is dummy variable that takes on the value one if it falls in the alcohol tobacco or gambling industry and is zero otherwise. Sindum1 is dummy variable that takes on the value one if it falls in the alcohol tobacco or gambling industry and is zero otherwise. ONEDIGDUM is a dummy variable which takes on the value one if the one digit sic code coincides with the one digit sic code of the alcohol, tobacco or gambling industry (codes 2 and 7), and is zero otherwise. LOGSIZE is the measure for size and is calculated by taking the natural logarithm of market capitalization. Market capitalization is calculated as end of year price on year t for firm i times shares outstanding on end of year t for firm i. LOGMB is the natural logarithm of the market to book value of firm i for year t and is measured as the end of year market capitalization divided by the end of year book value. The variable SP500 is a dummy variable that takes on the value one if the stock is listed on the S&P 500 index and zero otherwise. The rather similar variable NASD is a dummy variable that takes on the value one if the stock is listed on the NASDAQ and zero otherwise. STD is the standard deviation of daily returns for firm i during year t. BETA is the year-end beta reported by Compustat in year t for firm i. PRINV is the inverse of the firm i share price at the end of year t. RET is the average monthly return on stock i during year t. The numbers in brackets are the standard errors that are robust to heteroscedasticity and are clustered within firm over time. *** 1% level of significance; ** 5% level of significance; and * 10% level of significance.

Panel A: Regression analysis

	More constrained	**Less constrained**	**More constrained**	**Less constrained**
Sindum1	-.0149**	-.0355***	-.0161**	-.0392***
	(.0065)	(.0094)	(.0065)	(.0088)
Onedig-	.0036*	-.0022	.0053***	.0019
dum	(.0021)	(.0025)	(.0020)	(.0023)
Logsize	.0018***	-.0011	.0077***	.0100***
('000)	(.0006)	(.0012)	(.0007)	(.0012)
Beta	.0142***	.0434***	.0152***	.0421***
	(.0018)	(.0017)	(.0018)	(.0017)
LogMB	-.0068**	-.0038**	-.0099***	-.0112***
	(.0027)	(.0018)	(.0027)	(.0019)
Prinv	-.0030	-.0082**	-.0008	-.0033
	(.0018)	(.0041)	(.0014)	(.0028)
STD	-.6563***	-1.502***	-.5428***	-1.259***
	(.0509)	(.1033)	(.0540)	(.0985)
RET	.0693**	.1230***	.0420	.1684***
	(.0273)	(.0128)	(.0258)	(.0124)

NASD	-.0033**	.0149***	.0039***	.0285***
	(.0013)	(.0023)	(.0012)	(.0021)
SP500	.0806***	.0459***	.0672***	.0230***
	(.0034)	(.0043)	(.0035)	(.0043)
Year FE			yes	yes
R2	0.1249	0.1035	0.1776	0.3502
#obs	79,826	79,826	79,826	79,826

Panel B: Chi-square test analysis of difference of coefficients

Variables combination	**Difference (column 1 and 2)**
Sindum1 (more norm-constrained) – Sindum1 (less norm-constrained)	0.0206**

Table 9 presents the results of the righteous and ESG sin stock counterpart of the previous regression. Columns 1 and 2 show that both more and less norm-constrained institutional investors have a negative coefficient for sindum2. The significance level differs however. The more norm-constrained Sindum2 coefficient is significant at the 10% level while its less norm-constrained counterpart is significant at the 1% level. The coefficient for Sindum2 is -0.0049 for more norm-constrained investors and the coefficient for less norm-constrained investors is -0.0255. This means that more norm-constrained investors hold a relatively larger proportion of sin stocks than less norm-constrained investors. The chi-square test in panel B shows that these two coefficients are significantly different from each other at the 1% level. The difference amounts to 0.0206. Adding year fixed effects to the regression brings about a similar general finding (see column 3 and 4). The coefficient for Sindum2 is -0.0029 for more norm-constrained investors and the coefficient for less norm-constrained investors is -0.0095. The difference between these coefficients amounts to 0.0067 and is significant at the 5% level. Adding year fixed effects makes the difference between the two coefficients smaller. The relative proportion of sin stock shares held by less norm-constrained institutional investors is lower compared to the proportion of sin stock shares held by more norm-constrained institutional investors. In other words, the relationship between sin stock and institutional ownership is more negative for the less norm-constrained institutional investors than for the more norm-constrained institutional investor. Hence, the results in this section are not in line with the hypothesis. The general results in this section are similar to the results of the first sin stock selection method and therefore also oppose the social norm hypothesis.

The coefficient of the righteous stock dummy, which also can be found in table 9, is -0.0112 for more norm-constrained investors and is 0.0049 for less norm-constrained investors. The former is significant

at 1% level and the latter is not significant at the 10% level. The difference between these coefficients is -0.0161 and is significant at the 1% level (see table 9 panel B). The interpretation behind this is that more norm-constrained investors hold relatively less righteous stock than less norm-constrained investors. This is contradictory to the hypothesis. Adding year fixed effects, changes the coefficient for more norm-constrained investors to -0.0053 and to -0.0166 for less norm-constrained investors. Both relationships are significant at the 1% level. This finding is contradictory to the previous finding because now less norm-constrained investors hold relatively less righteous stock than more norm-constrained investors. The results in panel B show that the difference between the coefficients is 0.0113 and that the difference is significant at the 1% level. The second finding is in line with the hypothesis that more norm-constrained institutional investors hold a relatively larger proportion of righteous stock than less norm-constrained institutional investors. In conclusion, the results regarding righteous stocks provide mixed evidence.

Table 9
Sin and righteous stock ownership of more and less norm-constrained institutional investors

This table presents a pooled regression including two dependent variables and control variables for years 1990 to 2016. The total sample consists of 24,975 observations. Proportion institutional ownership for more norm-constrained investors is calculated by summing up the total shares owned by shareholder types 1, 2 and 5 for year t and dividing by total number of shares outstanding for year t. Proportion institutional ownership for less norm-constrained investors is calculated by summing up the total shares owned by shareholder types 3 and 4 for year t and dividing by total number of shares outstanding for year t. Rightdum is a dummy variable that takes on the value 1 if it is denoted as 1 on any of the positive counterpart MSCI KLD ESG factors human rights, child labor, waste management and climate change. LOGSIZE is the measure for size and is calculated by taking the natural logarithm of market capitalization. Market capitalization is calculated as end of year price on year t for firm i times shares outstanding on end of year t for firm i. LOGMB is the natural logarithm of the market to book value of firm i for year t and is measured as the end of year market capitalization divided by the end of year book value. The variable SP500 is a dummy variable that takes on the value one if the stock is listed on the S&P 500 index and zero otherwise. The rather similar variable NASD is a dummy variable that takes on the value one if the stock is listed on the NASDAQ and zero otherwise. STD is the standard deviation of daily returns for firm i during year t. BETA is the year-end beta reported by Compustat in year t for firm i. PRINV is the inverse of the firm i share price at the end of year t. RET is the average monthly return on stock i during year t. The numbers in brackets are the standard errors that are robust to heteroscedasticity and are clustered within firm over time. *** 1% level of significance; ** 5% level of significance; and * 10% level of significance.

Panel A: Regression analysis				
	More constrained	**Less constrained**	**More constrained**	**Less constrained**
Sindum2	-.0049*	-.0255***	-.0029	-.0095***
	(.0026)	(.0043)	(.0026)	(.0035)
Rightdum	-.0112***	.0049	-.0053***	-.0166***
	(.0022)	(.0041)	(.0019)	(.0031)
Logsize	.0013	.0068***	.0046***	.0013
('000)	(.0008)	(.0016)	(.0008)	(.0014)
Beta	.0045***	.0139***	.0094***	.0192***
	(.0012)	(.0022)	(.0013)	(.0020)
LogMB	-.0049***	-.0024	-.0050***	.0017
	(.0009)	(.0018)	(.0009)	(.0016)
Prinv	-.0347***	-.0764***	-.0259***	-.0901***
	(.0060)	(.0134)	(.0048)	(.0152)
STD	-.6056***	.2724***	-.7990***	-.8998***
	(.0498)	(.1049)	(.0779)	(.1269)
RET	.0417***	-.1364***	.0417***	.0138
	(.0106)	(.0229)	(.0135)	(.0223)
NASD	-.0092***	.0057**	-.0038**	.0060**
	(.0015)	(.0029)	(.0015)	(.0027)
SP500	.0467***	-.0048	.0278***	-.0021
	(.0026)	(.0047)	(.0025)	(.0038)
Year FE			yes	yes
R2	0.1904	0.0231	0.3269	0.4428
#obs	24,972	24,972	24,972	24,972

Panel B: Chi-square test analysis of difference of coefficients	
Variables combination	**Difference (column 1 and 2)**
Sindum2 (more norm-constrained) – Sindum2 (less norm-constrained)	0.0206***
Rightdum (more norm-constrained) - Rightdum (less norm-constrained)	-0.0161***
Variables combination	**Difference (column 3 and 4)**
Sindum2 (more norm-constrained) – Sindum2 (less norm-constrained)	0.0067**

Rightdum (more norm-constrained) - Rightdum (less norm-constrained)	0.0113***

Hong and Kacperczyk (2009) encountered the same issue regarding the measurement error in shareholder type 5 and restricted their data to before 1998. However, they also estimated the results after 1998 for completeness and found similar results. This does not hold for the results in this paper. Excluding shareholder type 5 from the data after 1997 has a large implication on the results. I estimated the results for hypothesis 4 again with the 'all other' shareholde type. These results are displayed in appendix tables 14 and 15. Regarding Sindum1, I find clear eviddence in favor of the social norm hypothesis. The coefficient for more norm-constrained institutional investors (shareholder types 1,2 and 5) is -0.0985 and the coefficient for less norm-constrained institutional investors is -0.0350. Both coefficients are significant at the 1% level and are also significantly different at the 1% level (see panel B). Similar results are found when adding year fixed effects. The results in appendix table 14 accept hypothesis 4A because the coefficient for more norm-constrained investors is significantly more negative than for less norm-constrained investors. These results provide additional evidence in favor of the social norm hypothesis.

Appendix table 15 presents the results of the righteous and sin stock counterpart of the previous regression. Regarding Sindum2, more norm-constrained investors have a relatively larger proportion of sin stocks than less norm-constrained investors (see column 1 and 2). This finding is not in line with the hypothesis, because of the less negative coefficient for more norm-constrained investors than the coefficient for less norm-constrained investors. However, adding year fixed effects to the regression makes a change to the results (see column 3 and 4). The coefficient for Sindum2 is signifcantly more negative for more norm-constrained investors than for less norm-constrained investors. In contrast to the previous finding this result is in line with the hypothesis that the relative proportion of sin stock shares held by more norm-constrained institutional investors is lower compared to the proportion of sin stock shares held by less norm-constrained institutional investors. This result is similar to the results of the first sin stock selection method. Sindum2 therefore provides partial support for the social norm hypothesis.

Appendix table 15 shows that for both tests, with and without year fixed effects, the righteous stock dummy coefficient for more norm-constrained investors is more negative than the coefficient for less norm-constrained investors. Both findings, with and without year fixed effects, are therefore not in line with the hypothesis that more norm-constrained institutional investors hold a larger proportion of righteous stock than less norm-constrained institutional investors.

All in all, the results on hypothesis 4 provide mixed evidence on the social norm hypothesis. As shown, adding shareholder type 5 has a strong effect on the results. The findings are further discussed in the discussion section.

6. Robustness checks

The next section contains a number of robustness checks. The main findings of the robustness checks are displayed in table 10. The output of the table is focused on the sin stock and righteous stock dummies because these variables form the backbone of this paper and the hypotheses. The remainder of the output of the robustness checks is not displayed in this paper for brevity reasons.

The first test checks the robustness of the results by changing the definition of the traditional sin stocks. This is done by adding the weapons industry as an additional sin industry. Adding the weapons industry to the sin stock dummy changes the coefficient for Sindum1 to roughly -0.1067, decreasing the effect by roughly 0.01 percentage points (see robustness check 1 in table 10). This means that the sin stock dummy effect is less pronounced for the weapons industry. The relationship remains highly significant at the 1% level. Adding this industry does not make a significant change to the general findings. Similarly the behavior of the sin stock dummy continues to follow the same pattern. Including the weapons industry also does not affect the third results section whereby the differences between sin industries are of interest. The coefficient for the industries remain rather similar and the significant differences remain. Adding the weapons industry as an additional sin industry also does not significantly change the general findings of the more and less norm-constrained results section.

Due to the construction of Sindum2 and Rightdum there are a number of observations/stocks that are denoted as righteous stock and also as sin stock. This happens when for example a firm undertakes exceptional human rights initiatives but also has severe controversies regarding non-GHG emissions. The coefficients for Sindum2 and Rightdum change to -0.0264 and -0.0656 respectively, whereby the former is significant at the 5% level and the latter at the 1% level (see robustness check 2 in table 10). Therefore, dropping observations that are considered both righteous and sin does not make a significant change to the findings. Similarly it does not make a significant difference to the rest of the results.

Another potential downfall regarding Sindum2 and Rightdum is that the factors that make up these two variables are rather specific. Therefore I execute a robustness check whereby I replace the existing factors human rights, child labor, waste management and climate change with other environmental

and social factors[14]. I create different subsets of ESG factors and create multiple sin and righteous stock dummies. Creating different sets of ESG factors and repeating the regressions analyses does not lead to a change in the general findings. The coefficients for the dummies of sin and righteous stock remain significant and of roughly the same magnitude. Robustness check 3 in table 10 shows an example whereby solely environmental factors are used to construct Sindum2 and Rightdum[15]. The differences between more and less norm–constrained investors also remain when changing the Sindum2 and Rightdum definitions.

Similarly I estimate the results for hypothesis one again using the MSCI KLD ESG factors alcohol involvement, gambling involvement and tobacco involvement. The effect of this adaptation on Sindum is shown in robustness check 4 in table 10. The robustness check is estimated twice. Once using the Sindum1 methodology and once using the Sindum2 methodology. The only differences between the two are the inclusion of the control variable Onedigdum and the inclusion of Rightdum in the regression. The Sindum coefficients remain negative and statistically significant and amount to roughly -0.04. Hence, the general conclusion that the proportion of institutional ownership of sin stocks is lower than the proportion of institutional ownership of non-sin stocks still holds.

Another robustness check that is performed in the literature is dropping all firms with book values below $10 million. The number of observations that have book values below $10 million is 6,845. Dropping these observations makes the sin stock dummy slightly more negative (decrease of 0.046 percentage points). See robustness check 5 in table 10. This means that the sin stock dummy effect is slightly more pronounced in larger stock. The conclusion for this part still holds. The screen makes the Sindum2 and Rightdum coefficients slightly more positive. These changes are negligible however and the general findings therefore remain the same. Regarding the second hypothesis adding the screen does not significantly affect the observed pattern of the sin stock dummy. Similarly it does not affect the general results of hypotheses three and four.

When constructing the control variable market to book value the logarithm is taken of the market to book value (M/B). However, when the book value is negative then the M/B value is negative. It is therefore not possible to construct the logarithm of M/B. Thus far in the paper these observations were dropped. For this robustness check I adjust the values of these observations. It is most logical to set the book values to zero. However this leads the value of the M/B variable to infinity. Therefore I

[14] The additional factors are Regulatory Compliance, Operational Waste, Environmental Opportunities - Green Buildings, Environmental Opportunities in Renewable Energy, Employee Involvement, Employee Relations, Women and Minority Contracting, Workforce Diversity and Community Impact. For a full description of the factors see the MSCI ESG KLD manual.

[15] Only one example is shown in table 10 because there are numerous combinations that can make up Sindum2 and Rightdum. This means that there are also numerous regression coefficients for Sindum2 and Rightdum.

set the M/B value and therefore also the Log M/B value at the maximum of the sample. This means that these observations are added to the right bottom of the distribution. The main point of this robustness check is not whether the Log M/B values are placed exactly on the right place in the distribution. The importance of adding these observations is that they are taken into account for the regressions, because these observations potentially have something in common that affects the proportion of institutional ownership. When estimating the regressions again including the observations with negative book values I find the following. The Sindum1 coefficient changes to -0.1138 and the coefficients for Sindum2 and Rightdum change to -0.0236 and -0.0652 respectively (see robustness check 6 in table 10). These values are comparable to the original values. Adding these extra observations with a negative book value also does not change the general findings regarding the rest of the hypotheses.

The standard errors are clustered at the firm level in this paper. Another method to cluster standard errors is by clustering at the industry level (see for example Hong and Kacperczyk (2009)). I estimate the regressions again using the Fama and French 49 industry groupings to cluster at the industry level. The results of robustness check 7 in table 10 show that the coefficient for Sindum1 remains similar in size and remains significant at the 1% level. The standard errors goes up slightly to 0.0258 (increase of 0.0008). The coefficient for Sindum2 also remains similar in size but loses some significance. However the relation is still significant at the 10% level. The standard error increases to 0.0153. The Rightdum coefficient keeps its 1% significance level and shows an increase of the standard error of 0.026. In conclusion, the general findings remain similar for all results after clustering standard errors at the industry level.

Furthermore, I check the robustness of my results by estimating my results again using newey-west standard errors. This robustness check is executed to test whether autocorrelation is an issue in the data. I estimate the results again using different numbers of lags. The effect of newey-west standard errors on Sindum and Rightdum is shown in table 10 robustness check 8. The test shows that all three coefficients remain similar as to the original value and the three coefficients remain significant at the 1% level. The standard errors however are smaller than the original standard errors. Furthermore all the results except for the findings for hypothesis 3 remain generally similar. The coefficient for the tobacco industry becomes significant at the 5% level when estimating the regression for hypothesis 3 using newey-west standard errors. The results also show that there is a significant difference between the alcohol and gambling industry regarding their relationship to the proportion of institutional ownership. The newey-west standard error method does not cluster at the firm-level and therefore detracts more statistical information from the individual observations of one firm. The results regarding the tobacco industry are therefore not surprising because of the small number of firms in

the tobacco industry. The findings that use firm-level clustered errors are likely to be the most conservative. I therefore stick to these standard errors for the conclusion of this paper. This robustness check does not make the findings of this paper less significant (only more). Hence the findings are robust with regards to newey-west standard errors.

Table 10
Summary of robustness checks

This table presents the summary of the robustness checks. The dependent variable is proportion of institutional ownership. This table consists of two regressions per row. One regression including Sindum1 and one regression including Sindum2 and Rightdum. The control variables for the Sindum 1 regression are Onedigdum, Logsize, Beta, LogMB, Prinv, STD, RET, NASD and SP500. The Sindum2 and Rightdum regression contains the same conrol variables except for Onedigdum. The total sample differs per robustness check and is displayed in the columns with the heading #obs. The numbers in brackets are the standard errors that are robust to heteroscedasticity and are clustered within firm over time, unless stated otherwise. The regressions are estimated without year fixed effects. Robustness check 8 regarding newey-west standard errors is executed using 3 lags. *** 1% level of significance; ** 5% level of significance; and * 10% level of significance.

	Sindum1	**R2**	**#obs**	**Sindum2**	**Rightdum**	**R2**	**#obs**
Original values	-.1183***	0.3126	81,662	-.0279***	-.0660***	0.0768	24,983
	(.0250)			(.0101)	(.0075)		
Robustness check 1	-.1067***	0.3124	81,662	NA	NA	NA	NA
(Weapons industry)	(.0244)						
Robustness check 2	NA	NA	NA	-.0264**	-.0656***	0.0775	24,312
(Right- and Sindum)				(.0118)	(.0084)		
Robustness check 3	NA	NA	NA	-.0323***	-.0624***	0.0775	24,721
(Other ESG factors)				(.0087)	(.0079)		
Robustness check 4	-.0386**	0.0680	25,578	-.0401**	-.0683***	0.0770	24,984
(MSCI alc, tob, gam)	(.0163)			(.0168)	(.0076)		
Robustness check 5	-.1229***	0.2501	74,817	-.0271***	-.0654***	0.0727	24,893
(B.V. under $10 mil)	(.0250)			(.0101)	(.0075)		
Robustness check 6	-.1138***	0.3077	84,896	-.0236**	-.0652***	0.0728	25,743
(Log M/B)	(.0251)			(.0101)	(.0075)		
Robustness check 7	-.1183***	0.3126	81,662	-.0279*	-.0660***	0.0768	24,983
(Industry level cluster)	(.0258)			(.0153)	(.0101)		
Robustness check 8	-.1183***	NA	81,662	-.0279***	-.0660***	NA	24,983
(Newey-West std. errors)	(.0140)			(.0071)	(.0061)		

7. Discussion

In this chapter I will restate my main findings, explain the meaning of those findings further, then relate the findings to existing literature and lastly I will aim to explain any unexpected or inconclusive results. The results of the first regression analysis whereby the traditional sin stock selection process was used showed that sin stocks have a significantly lower proportion of institutional ownership than non-sin stocks. This difference amounted to roughly 11 percentage points. Hong and Kacperczyk (2009) found a similar sign of the sin stock dummy. However, the results in this paper show a larger magnitude of the sin dummy effect. Hong and Kacperczyk (2009) found a difference of roughly 4.2 to 6.7 percentage points. The difference in results can potentially be explained by the different time periods in the datasets. I used the timespan 1990 to 2016 whereas Hong and Kacperczyk (2009) used 1980 to 2003. The results of the second results section support this explanation. This section showed that the effect of the sin dummy on institutional ownership has become more negative over time. This means that sin stocks have received a lower and lower proportion of institutional ownership over time. In the first five years of the data, there is no significant difference between sin stocks and non-sin stocks. However in the final 7 years of the data the difference amounts to roughly 19 percentage points. The negative increase of this effect was hypothesized and could be explained by the increase in popularity of SRI and CSR, which goes hand in hand with shunning of sin stocks by investors. When more institutional investors shun sin stocks, then institutional ownership of these sin stocks will decrease. Hong and Kacperczyk (2009) focussed on the Sindum coefficient for different time periods up to 2006. Their Sindum coefficient for the time period 1990 to 1994 is slightly more negative (-0.0305 vs. 0.0198) and their coefficients for the time periods 1995 to 1999 and 2000 to 2006 are slightly more positive (respectively -0.0513 vs -0.0919 and -0.0638 vs. -0.0854). It remains unclear why these coefficients differ. Similar research as done for hypothesis 2 has not been performed after 2006 according to the best of my knowledge so I am unable to link these results to existing literature.

Another finding of this paper is that sin stocks, when selected using MSCI KLD ESG factors, receive roughly 3 to 4 percentage points less institutional ownership than non-sin stocks. This is in line with the finding described above and therefore also in line with the finding of Hong and Kacperczyk (2009). One potential explanation for the difference in magnitude of the sindum dummy coefficients could be that institutional investors include mainly traditional sin stocks in their exclusion lists and not the MSCI KLD ESG sin factors. This would lead to a difference in proportion of institutional ownership between the two types of sin stocks. The definite answer that explains the difference remains unknown. The question that remains is as follows. What explains the difference in magnitude of the sin stock dummy between traditional sin stocks and other ESG sin stocks? This is a question that is potentially interesting and is therefore recommended for future research.

This paper introduced the new term righteous stock and a negative relationship between the righteous stock dummy and institutional ownership was found. The results ended up with a negative coefficient of roughly 7 percentage points. This means that righteous stocks have a 7 percentage points lower institutional ownership in comparison to non-righteous stock. This result is in line with Fernando et al. (2010) who found that positive ("green") environmental performers have lower institutional ownership. The results are not in line with Graves & Waddock (1994) because in their research social performancers have a positive but insignificant relationship with the percentage of shares held by institutions. It is important to be aware of the differences between the measures for selecting positive stocks. Fernando et al. (2010) merely focus on an environmental factor with just over 7000 observations and Graves & Waddock (1994) solely focus on social factors with a dataset of 430 firms from the S&P500. This paper constructs a more complete measure (righteous stocks) that includes both social and environmental factors with an extensive dataset of nearly 25.000 observations. One potential explanation for finding a negative coefficient could be that some institutional investors believe that investments in such righteous stocks diminishes shareholder value. Righteous stock have a tendency to value multiple sorts of welfare such as community welfare, employee welfare and trading partner welfare. Institutional investors who purely focus on economic welfare might view these investments as a waste or destruction of shareholder value. These institutional investors therefore stay away from these stocks, which leads to lower institutional ownership.

The next main finding of this paper is that the three traditional sin stock industries do not have significantly different relationships with institutional ownership. Sin stocks in the tobacco or alcohol industry experience roughly 15 percentage points lower institutional ownership than non-sin stocks. However, the relationship for the tobacco industry is not significant at the 10% level. The explanation for this lack of significance could be the sample size of the tobacco industry. The tobacco industry includes 2 to 5 tobacco stocks per year which is relatively small. Therefore, the results of this industry might not represent the majority of this industry. Sin stocks in the gambling industry experience roughly 8 percentage points lower institutional ownership than non-sin stocks. These findings suggest that institutional investors treat the gambling industry different than the tobacco or alcohol industry. One potential explanation for this is that investors view the gambling industry as less sinful or harmful than the alcohol industry. This heterogeneity among investors regarding their view on sinful leads to differences in their exlusion lists, which is followed by an appurtenant effect on institutional ownership. However, the coefficients for the different sin industries are not significantly different from each other. Therefore I cannot draw the conclusion that the gambling industry is shunned to a lesser degree. The explanation that follows then is that institutional investors do not view particular sin industries as more or less sinful. This then leads to a statistically similar effect on the proportion of

institutional ownership for the different sin industries. Similar research as done for hypothesis 3 has not been performed according to the best of my knowledge so I am unable to link these results to existing literature.

The final results of this paper make a distinction between more and less norm-constrained institutional investors. I find that the sin stock dummy effect, where the traditional method is used, for more norm-constrained investors is less negative than for less norm-constrained investors. These results provide evidence opposing the social norm hypothesis and therefore also oppose the findings of Hong and Kacperczyk (2009). One potential explanation for the difference in findings is the exclusion of shareholder type 5. The results in appendix table 14 show that when including this shareholder type then the findings are similar to what Hong and Kacperczyk (2009) found in their research. The general finding is then similar, however there is a difference in magnitude. Their difference amounts to 1% whereas the difference in appendix table 14 amounts to 5%. This could be explained by the larger negative effect of the sin stock dummy in later years that is found in hypothesis two. Selecting sin stocks the 'ESG way' brings about similar results. The results also show that less norm-constrained institutional investors hold a relatively lower proportion of sin stocks than more norm-constrained institutional investors. This subsection therefore also opposes the finding of Hong and Kacperczyk (2009) and the social norm hypothesis.

The righteous stock counterpart of this section exhibits unexpected results. The first model without year fixed effects shows that more norm-constrained investors hold a relatively lower proportion of righteous stocks than less norm-constrained investors. The second model with year fixed effects contains a contradictory finding, which is in line with the hypothesis. The negative coefficient for the righteous stock dummy in 3 out of 4 relationships is in line with Fernando et al. (2010). Fernando et al. (2010) found that all institutional types except other institutions (including universities, pension plans and employee stock ownership plans) hold significantly lower fractions of the shares of green firms.

The measurement error in shareholder type 5 described by Gompers and Metrick (2001) provides a limitation to this study. Eliminating this measurement error caused the deletion of data on shareholder type 5 which also included valuable and correct data on for example pension fund holdings. I in fact compared the difference between banks (type 1) and insurance companies (type 2) and investment companies and their managers (type 3) and independent investment advisor (type 4) (1 and 2 vs. 3 and 4) after 1997, whereas the prefered comparison is between shareholder types 1, 2 and 5 vs. 3 and 4. The results in appendix tables 14 and 15 showed that the results differ when including shareholder type 5. This therefore limits the plausibility of the findings in the main text which leads to the

conclusion that this paper provides mixed evidence on the difference between more and less norm-constrained institutional investors regarding their relation to institutional owenership and sin stocks.

8. Conclusion

In this paper I examined the relationship between institutional ownership and sin and righteous stocks using panel data for the years 1990 to 2016. The results showed that when selecting sin stocks in the traditional way, sin stocks have a significantly lower proportion of institutional ownership than non-sin stocks. Similarly, when selecting sin stock using different MSCI KLD ESG factors sin stocks also have a lower proportion of institutional ownership than non-sin stocks. Furthermore, the newly introduced term righteous stock had a negative relationship with institutional ownership. This research also presented the behavior of sin stock institutional ownership over time and found that the difference in proportion of institutional ownership between sin stocks and non-sin stocks has become larger over time. Furthermore, this paper also examined differences in institutional ownership for different sin industries and found that there is no significant difference between the three sin industries and their relationship with institutional ownership. Lastly, the hypothesis that more norm-constrained institutional investors hold a lower proportion of sin stocks than less norm-constrained investors cannot be accepted. The results regarding the differences between more and less norm-constrained institutional investors and their relationships to the proportion of sin and righteous stock ownership provide mixed evidence.

This paper contributes to the existing literature in the following ways. First of all this paper introduced a new way of selecting sin stocks and examined the sin stock and institutional ownership relationship using this new method. This adds value to the literature by providing more evidence on the robustness of the social norm hypothesis. Second of all, this paper introduced a new term: righteous stocks, and examined the relationship between righteous stock and institutional ownership. Evidence on the relationship between institutional ownership and sin and righteous contributes to the literature because it contains information on how institutional investors treat firms that act socially, ethically or economically wrong or right according to the norms of society. This paper shows that institutional investors in general 'punish' both sin activities and righteous activities by lower institutional ownership. This information is potentially useful for policy makers, owners of these firms or even new firms who are considering new business activities that contain aspects of sin or righteous factors. Third of all, this paper presents the behavior of the relationship between sin stocks and institutional ownership over time for recent years, which provides more information on how institutional investors treat sin stocks. Fourth of all, this paper adds value by zooming in on the relationship of sin stocks and

institutional ownership by examining different sin stock industries and their individual relationship to institutional ownership. This has not been done before according to the best of my knowledge and therefore adds to the completeness of the state of the academic literature (contribution for academics). The reason why this contributes for practitioners is because it provides information on whether investors distinguish between different sin industries. This paper shows that there is no significant difference between sin industries regarding their relation to institutional ownership. For instance, this is then valuable information for a sin industry firm looking to invest in another sin industry firm because the firm knows that is not likely to incur an additional institutional ownership penalty for this investment[16]. Finally, this paper contributes by testing the social norm hypothesis with more recent data, thereby testing whether social norms still have an effect on markets in recent years.

Although this research was carefully prepared, I am still aware of its limitations and shortcomings, beside the exclusion of shareholder type 5. The first limitation considers the sample size of the tobacco industry. Even though the total sample size is substantial the sample size of the tobacco industry includes 2 to 5 tobacco stocks per year. The results of this industry might therefore not represent the majority of this industry. The second sin stock measure and the righteous stock measure have their flaws. They are both dummy variables that take on the value one if one of the underlying factors has a one value. This is a rather black and white measure and it is therefore better to use a non-dichotomous dummy, or use a continues variable. These limitations provide new pathways for future research to improve the state of the academic literature.

The newly introduced concept of righteous stock is a potentially fruitful new research direction. The coefficients for righteous stock (Rightdum) are highly significant in this paper. However, it remains unclear why the relationship with institutional ownership is generally negative and also what the driver(s) of this significant negative relationship is (are). Hence, future research on this topic is recommended and required to shed more light on this relationship.

[16] This reasoning only holds with regards to this research. Incurring an institutional ownership penalty remains possible for several other reasons.

9. Reference list

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence. *The quarterly journal of economics, 94*(4), 749-775.

Barnea, A., & Rubin, A. (2010). Corporate social responsibility as a conflict between shareholders. *Journal of business ethics, 97*(1), 71-86.

Beal, D., & Goyen, M. (1998). 'Putting your money where your mouth is' a profile of ethical investors. *Financial Services Review, 7*(2), 129-143.

Blitz, D., & Fabozzi, F. J. (2017). Sin Stocks Revisited: Resolving the Sin Stock Anomaly. *The Journal of Portfolio Management, 44*(1), 105-111.

Blume, M. E., & Keim, D. B. (2014). The changing nature of institutional stock investing. *Critical Finance Review, 6*, 1-41.

Bollen, N. P. (2007). Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative Analysis, 42*(3), 683-708.

Brigham, G., Gapenski, L. C., & Ehrhardt. (1999). *Financial Management: Theory and Practice* (9th ed.). Dryden Press.

Capelle-Blancard, G., & Monjon, S. (2012). Trends in the literature on socially responsible investment: Looking for the keys under the lamppost. *Business Ethics: A European Review, 21*(3), 239-250.

Carroll, A. B. (1999). Corporate social responsibility: Evolution of a definitional construct. *Business & society, 38*(3), 268-295.

Coffey, B. S., & Fryxell, G. E. (1991). Institutional ownership of stock and dimensions of corporate social performance: An empirical examination. *Journal of Business Ethics, 10*(6), 437-444.

Coval, J. D., & Moskowitz, T. J. (1999). Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance, 54*(6), 2045-2073.

Derwall, J., Koedijk, K., & Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. *Journal of Banking & Finance, 35*(8), 2137-2147.

Drumwright, M. E. (1994). Socially responsible organizational buying: environmental concern as a noneconomic buying criterion. *The Journal of Marketing, 58*(3), 1-19.

Durand, R. B., Koh, S., & Tan, P. L. (2013). The price of sin in the Pacific-Basin. *Pacific-Basin Finance Journal, 21*(1), 899-913.

Fabozzi, F. J., Ma, K. C., & Oliphant, B. J. (2008). Sin stock returns. *Journal of Portfolio Management, 35*(1), 82-94.

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

Fama, E. F., & French, K. R. (1997). Industry costs of equity. *Journal of financial economics, 43*(2), 153-193.

Fauver, L., & McDonald IV, M. B. (2014). International variation in sin stocks and its effects on equity valuation. *Journal of corporate finance, 25*, 173-187.

Fernando, C., Sharfman, M., & Uysal, V. (2010). Does greenness matter? The effect of corporate environmental performance on ownership structure, analyst coverage and firm value. *FMA European Conference Hamburg, Germany.*

Fombrun, C. (1995). *Reputation: Realizing value from the corporate image.* John Wiley & Sons, Ltd.

Gompers, P. A., & Metrick, A. (2001). Institutional investors and equity prices. *The quarterly journal of Economics, 116*(1), 229-259.

Goss, A., & Roberts, G. S. (2011). The impact of corporate social responsibility on the cost of bank loans. *Journal of Banking & Finance, 35*(7), 1794-1810.

Graves, S. B., & Waddock, S. A. (1994). Institutional owners and corporate social performance. *Academy of Management journal, 37*(4), 1034-1046.

Heinrichs, N., Rapee, R. M., Alden, L. A., Bögels, S., Hofmann, S. G., Oh, K. J., & Sakano, Y. (2006). Cultural differences in perceived social norms and social anxiety. *Behaviour Research and Therapy, 44*(8), 1187-1197.

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics, 93*(1), 15-36.

Johnson, R. A., & Greening, D. W. (1999). The effects of corporate governance and institutional ownership types on corporate social performance. *Academy of management journal, 42*(5), 564-576.

Krumsiek, B. (2003). Socially responsible high tech companies: Emerging issues. *Journal of Business Ethics, 43*(3), 179-187.

Kumar, A., Page, J. K., & Spalt, O. G. (2011). Religious beliefs, gambling attitudes, and financial market outcomes. *Journal of Financial Economics, 102*(3), 671-708.

Laufer, W. S. (2003). Social screening of investments: an introduction. *Journal of Business Ethics, 43*(3), 163-165.

Lobe, S., & Walkshäusl, C. (2016). Vice versus virtue investing around the world. *Review of Managerial Science, 10*(2), 303-344.

MSCI Launches New Factor ESG Target Indexes . (2017, September 28). Retrieved from MSCI: https://www.msci.com/documents/10199/3271eeb8-9dfa-4a92-a1bc-3d6f3feb4592

Neubaum, D. O., & Zahra, S. A. (2006). Institutional ownership and corporate social performance: The moderating effects of investment horizon, activism, and coordination. *Journal of Management, 32*(2), 108-131.

Poschke, M. (2011). *The firm size distribution across countries and skill-biased change in entrepreneurial technology*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2403128

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance, 32*(9), 1723-1742.

Robin, D. P., & Reidenbach, R. E. (1987). Social responsibility, ethics, and marketing strategy: Closing the gap between concept and application. *The Journal of Marketing, 51*(1), 44-58.

Roca, E., Wong, V. S., & Anand Tularam, G. (2010). Are socially responsible investment markets worldwide integrated? *Accounting Research Journal, 23*(3), 281-301.

Salaber, J. M. (2007). *The determinants of sin stock returns: Evidence on the European market.* Retrieved from https://ssrn.com/abstract=1071746

Sethi, S. P. (2005). Investing in socially responsible companies is a must for public pension funds–because there is no better alternative. *Journal of Business Ethics, 56*(2), 99-129.

Skypala, P. (2017, 11 27). ESG investing and smart beta combination grows in popularity. *Financial times*.

Sparkes, R., & Cowton, C. J. (2004). The maturing of socially responsible investment: A review of the developing link with corporate social responsibility. *Journal of Business Ethics, 52*(1), 45-57.

Statman, M. (2005). *The religions of social responsibility*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=774386

Statman, M., & Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal, 65*(4), 33-46.

Stulz, R. M., & Williamson, R. (2003). Culture, openness, and finance. *Journal of financial Economics, 70*(3), 313-349.

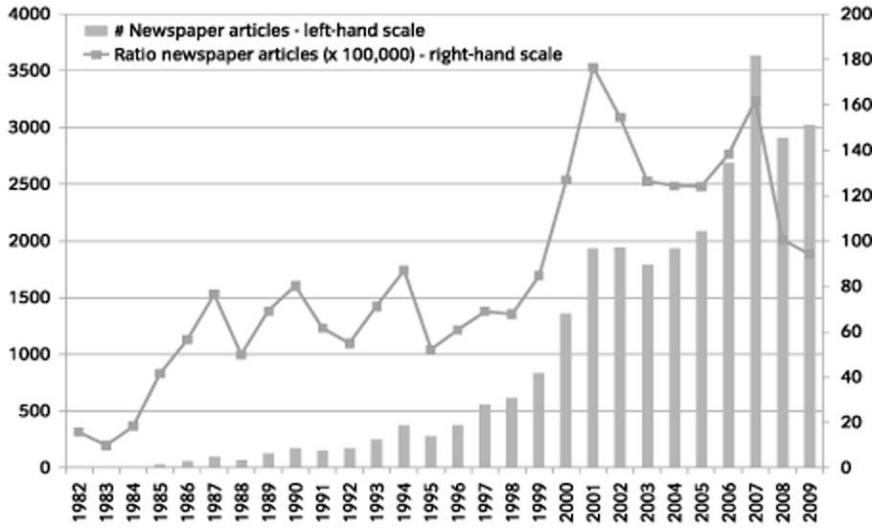
Zyglidopoulos, S. C., Georgiadis, A. P., Carroll, C. E., & Siegel, D. S. (2012). Does media attention drive corporate social responsibility? *Journal of Business Research, 65*(11), 1622-1627.

10. Appendix

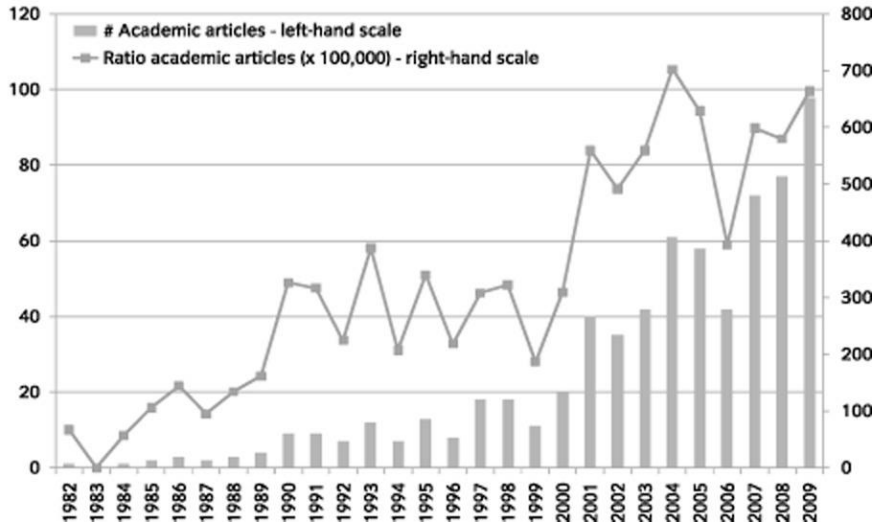
Appendix graph 2

The popularity of SRI in (a) newspapers and (b) academic journals.

(a) Newspapers



(b) Academic journals



Source: (Capelle-Blancard & Monjon, 2012)

Appendix table 11
Correlation matrix of all variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Institutional ownership (%)	1.000																
2. Sindum1	-0.029	1.000															
3. Alcdum	-0.025	0.615	1.000														
4. Gamdum	-0.017	0.724	-0.005	1.000													
5. Tobdum	-0.001	0.298	-0.002	-0.002	1.000												
6. Sindum2	0.138	-0.007	-0.005	-0.013	0.017	1.000											
7. Rightdum	0.133	0.015	0.019	-0.006	0.024	0.265	1.000										
8. Ffdum	0.026	-0.194	-0.129	-0.132	-0.059	-0.001	-0.000	1.000									
9. Onedigdum	-0.000	0.034	0.097	-0.054	0.047	-0.033	-0.010	-0.259	1.000								
10. Logsize	0.472	0.037	0.032	0.001	0.051	0.208	0.266	0.004	0.025	1.000							
11. Beta	0.304	-0.015	-0.029	0.009	-0.013	0.079	0.032	0.049	-0.005	0.318	1.000						
12. LogMB	0.095	0.026	0.022	-0.001	0.046	0.001	0.039	-0.059	0.137	0.410	0.162	1.000					
13. STD	-0.351	-0.025	-0.023	-0.004	-0.026	-0.089	-0.126	0.033	0.033	-0.520	0.006	-0.113	1.000				
14. Ret	0.047	0.001	-0.002	0.004	0.001	-0.006	-0.008	-0.002	0.003	0.116	0.069	0.291	0.032	1.000			
15. Prinv	-0.213	-0.003	0.000	-0.001	-0.007	-0.035	-0.040	0.005	-0.000	-0.320	-0.102	-0.174	0.405	-0.132	1. 000		
16. NASD	0.041	-0.003	-0.007	0.012	-0.023	-0.116	-0.101	0.038	0.087	-0.173	-0.025	0.050	0.109	0.054	-0.054	1.000	
17. SP500	0.275	0.025	0.035	-0.015	0.049	0.255	0.367	-0.012	0.003	0.480	0.075	0.121	-0.221	0.009	-0.079	-0.166	1.000

Appendix table 12
Number of sin and righteous stocks

This table presents the number of firms denoted as a '1' for the variables Waste management (env_str_b), climate change (env_str_d), Human Rights (hum_str_x), Waste management (env_con_d), Climate Change (env_con_f), Human Rights (hum_con_k) and Child labor (emp_con_g) by year for years 1995 to 2016. Total N is the total number of '1's' for the factor of the column.

Year	Waste Management (+)	Climate Change (+)	Human Rights (+)	Waste Management (-)	Climate Change (-)	Human Rights (-)	Child labor (-)
1995	14	24	3	17	n/a	n/a	n/a
1996	19	25	11	26	n/a	n/a	n/a
1997	24	30	5	27	n/a	n/a	n/a
1998	30	31	1	32	n/a	n/a	n/a
1999	23	33	2	26	27	n/a	n/a
2000	16	33	2	25	28	n/a	n/a
2001	15	42	2	62	48	n/a	n/a
2002	24	37	0	43	55	n/a	n/a
2003	23	56	0	47	83	n/a	n/a
2004	20	37	1	89	85	n/a	n/a
2005	22	37	2	67	79	n/a	n/a
2006	23	49	3	82	83	n/a	n/a
2007	20	69	5	82	93	n/a	n/a
2008	17	82	6	80	105	n/a	n/a
2009	16	84	6	85	100	n/a	n/a
2010	58	256	25	37	99	n/a	n/a
2011	59	235	29	39	90	n/a	n/a
2012	37	135	39	89	14	12	8
2013	72	158	74	142	32	9	11
2014	52	139	42	111	13	8	8
2015	77	167	50	95	12	6	8
2016	72	184	25	25	1	5	13
Total N	733	1943	333	1328	1047	40	48

Appendix table 13
Sin stock institutional ownership per year

This table contains an overview of the regression coefficients of the variable Sindum per year for the years 1990 to 2016. The dependent variable is institutional ownership and the control variables that are included is in this regression are: Onedigdum, LOGSIZE, BETA, LOGMB, STD, RET, PRINV, SP500 and NASD. Percentage institutional ownership is defined as shares owned of stock i by institutional investors for year t divided by the total number of shares outstanding of stock i for year t. The variable of interest is Sindum1. Sindum1 is dummy variable that takes on the value one if it falls in the alcohol tobacco or gambling industry and is zero otherwise. The total sample consists of 81,662 observations. The standard errors are robust to heteroscedasticity and are clustered within firm over time. *** 1% level of significance; ** 5% level of significance; and * 10% level of significance.

Year	Sindum1 coefficient	Standard error	#obs
1990	.0334	.0461	2,383
1991	.0491	.0468	2,317
1992	.0484	.0479	2,597
1993	-.0710**	.0355	2,789
1994	-.0706**	.0295	3,052
1995	-.1007***	.0251	3,282
1996	-.0986***	.0236	3,612
1997	-.0583**	.0288	3,750
1998	-.0933***	.0305	3,773
1999	-.0882***	.0325	3,602
2000	-.0879**	.0371	3,536
2001	-.0500	.0432	3,254
2002	-.0703*	.0416	3,147
2003	-.0902**	.0380	3,065
2004	-.0973**	.0410	3,083
2005	-.1115***	.0415	3,050
2006	-.0751*	.0388	3,047
2007	-.1538***	.0428	2,972
2008	-.1365***	.0457	2,882
2009	-.1713***	.0466	2,779
2010	-.1933***	.0394	2,727
2011	-.1956***	.0375	2,729
2012	-.1828***	.0446	2,738
2013	-.1909***	.0444	2,759
2014	-.1793***	.0416	2,886
2015	-.2140***	.0393	2,953
2016	-.1501***	.0444	2,898

Appendix table 14
Sin stock ownership of more and less norm-constrained institutional investors

This table presents a pooled regression including two dependent variables and control variables for years 1990 to 2016. The total sample consists of 80,851 observations. Proportion institutional ownership for more norm-constrained investors is calculated by summing up the total shares owned by shareholder types 1, 2 and 5 for year t and dividing by total number of shares outstanding for year t. Proportion institutional ownership for less norm-constrained investors is calculated by summing up the total shares owned by shareholder types 3 and 4 for year t and dividing by total number of shares outstanding for year t. Sindum1 is dummy variable that takes on the value one if it falls in the alcohol tobacco or gambling industry and is zero otherwise. Sindum1 is dummy variable that takes on the value one if it falls in the alcohol tobacco or gambling industry and is zero otherwise. ONEDIGDUM is a dummy variable which takes on the value one if the one digit sic code coincides with the one digit sic code of the alcohol, tobacco or gambling industry (codes 2 and 7), and is zero otherwise. LOGSIZE is the measure for size and is calculated by taking the natural logarithm of market capitalization. Market capitalization is calculated as end of year price on year t for firm i times shares outstanding on end of year t for firm i. LOGMB is the natural logarithm of the market to book value of firm i for year t and is measured as the end of year market capitalization divided by the end of year book value. The variable SP500 is a dummy variable that takes on the value one if the stock is listed on the S&P 500 index and zero otherwise. The rather similar variable NASD is a dummy variable that takes on the value one if the stock is listed on the NASDAQ and zero otherwise. STD is the standard deviation of daily returns for firm i during year t. BETA is the year-end beta reported by Compustat in year t for firm i. PRINV is the inverse of the firm i share price at the end of year t. RET is the average monthly return on stock i during year t. The numbers in brackets are the standard errors that are robust to heteroscedasticity and are clustered within firm over time. *** 1% level of significance; ** 5% level of significance; and * 10% level of significance.

Panel A: Regression analysis

	More constrained	Less constrained	More constrained	Less constrained
Sindum1	-.0985***	-.0350***	-.0918***	-.0389***
	(.0208)	(.0094)	(.0189)	(.0037)
Onedig-	.0200**	-.0024	.0125	.0017
dum	(.0093)	(.0024)	(.0090)	(.0023)
Logsize	.0492***	-.0004	.0238***	.0102***
('000)	(.0025)	(.0011)	(.0027)	(.0011)
Beta	.0733***	.0434***	.0645***	.0414***
	(.0085)	(.0016)	(.0086)	(.0016)
LogMB	-.0577***	-.0039**	-.0423***	-.0111***
	(.0139)	(.0017)	(.0140)	(.0018)
Prinv	-.0013	-.0069**	-.0143**	-.0027
	(.0041)	(.0032)	(.0059)	(.0021)
STD	-1.0252***	-1.4810***	-1.3965***	-1.1962***
	(.1333)	(.0992)	(.1802)	(.0906)
RET	.1329	.1215***	.2237**	.1627***
	(.1298)	(.0121)	(.1235)	(.0117)

NASD	.0810***	.0156***	.0474***	.0278***
	(.0051)	(.0022)	(.0048)	(.0020)
SP500	.0535***	.0453***	.1098***	.0233***
	(.0132)	(.0043)	(.0129)	(.0042)
Year FE			yes	yes
R2	0.1028	0.1082	0.1755	0.3554
#obs	80,851	80,851	80,851	80,851

Panel B: Chi-square test analysis of difference of coefficients

Variables combination	**Difference (column 1 and 2)**
Sindum1 (more norm-constrained) – Sindum1 (less norm-constrained)	-0.0635***

Appendix table 15
Sin and righteous stock ownership of more and less norm-constrained institutional investors

This table presents a pooled regression including two dependent variables and control variables for years 1990 to 2016. The total sample consists of 24,975 observations. Proportion institutional ownership for more norm-constrained investors is calculated by summing up the total shares owned by shareholder types 1, 2 and 5 for year t and dividing by total number of shares outstanding for year t. Proportion institutional ownership for less norm-constrained investors is calculated by summing up the total shares owned by shareholder types 3 and 4 for year t and dividing by total number of shares outstanding for year t. Rightdum is a dummy variable that takes on the value 1 if it is denoted as 1 on any of the positive counterpart MSCI KLD ESG factors human rights, child labor, waste management and climate change. LOGSIZE is the measure for size and is calculated by taking the natural logarithm of market capitalization. Market capitalization is calculated as end of year price on year t for firm i times shares outstanding on end of year t for firm i. LOGMB is the natural logarithm of the market to book value of firm i for year t and is measured as the end of year market capitalization divided by the end of year book value. The variable SP500 is a dummy variable that takes on the value one if the stock is listed on the S&P 500 index and zero otherwise. The rather similar variable NASD is a dummy variable that takes on the value one if the stock is listed on the NASDAQ and zero otherwise. STD is the standard deviation of daily returns for firm i during year t. BETA is the year-end beta reported by Compustat in year t for firm i. PRINV is the inverse of the firm i share price at the end of year t. RET is the average monthly return on stock i during year t. The numbers in brackets are the standard errors that are robust to heteroscedasticity and are clustered within firm over time. *** 1% level of significance; ** 5% level of significance; and * 10% level of significance.

Panel A: Regression analysis

	More constrained	**Less constrained**	**More constrained**	**Less constrained**
Sindum2	-.0061	-.0255***	-.0263**	-.0096***
	(.0118)	(.0043)	(.0119)	(.0035)
Rightdum	-.0646***	.0048	-.0591***	-.0166***
	(.0064)	(.0041)	(.0062)	(.0031)

Logsize	.0183***	.0067***	.0103***	.0012
('000)	(.0037)	(.0016)	(.0038)	(.0013)
Beta	.0601***	.0138***	.0474***	.0192***
	(.0051)	(.0021)	(.0050)	(.0020)
LogMB	-.0131***	-.0024	-.0151***	.0017
	(.0039)	(.0018)	(.0039)	(.0016)
Prinv	-.1946***	-.0763***	-.1891***	-.0902***
	(.0319)	(.0133)	(.0316)	(.0152)
STD	-1.688***	-.2722***	-1.855***	-.8993***
	(.2094)	(.1049)	(.2706)	(.1270)
RET	-.2072***	-.1364***	.0794*	.0128
	(.0416)	(.0228)	(.0450)	(.0223)
NASD	.0219***	.0057**	.0062	.0059**
	(.0062)	(.0029)	(.0061)	(.0026)
SP500	-.0470***	-.0047	(.0143	-.0021
	(.0094)	(.0047)	(.0093)	(.0037)
Year FE			yes	yes
R2	0.0610	0.0232	0.2032	0.4425
#obs	24,975	24,975	24,975	24,975

Panel B: Chi-square test analysis of difference of coefficients

Variables combination	**Difference (column 1 and 2)**
Sindum2 (more norm-constrained) – Sindum2 (less norm-constrained)	0.0194*
Rightdum (more norm-constrained) - Rightdum (less norm-constrained)	-0.0694***
Variables combination	**Difference (column 3 and 4)**
Rightdum (more norm-constrained) - Rightdum (less norm-constrained)	-0.0425***
Sindum2 (more norm-constrained) – Sindum2 (less norm-constrained)	-0.0167*