Journal of Financial Economics 93 (2009) 15–36

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

# The price of sin: The effects of social norms on markets[☆]

Harrison Hong[a,*], Marcin Kacperczyk[b]

[a] Princeton University, Princeton, NJ 08540, USA and NBER
[b] New York University, New York, NY 10012, USA and NBER

ARTICLE INFO

*Article history:*
Received 6 March 2007
Received in revised form
30 April 2008
Accepted 16 September 2008
Available online 8 April 2009

*JEL classification:*
G11
D71

*Keywords:*
Expected stock returns
Institutional ownership
Social norms
Sin stocks

ABSTRACT

We provide evidence for the effects of social norms on markets by studying "sin" stocks—publicly traded companies involved in producing alcohol, tobacco, and gaming. We hypothesize that there is a societal norm against funding operations that promote vice and that some investors, particularly institutions subject to norms, pay a financial cost in abstaining from these stocks. Consistent with this hypothesis, we find that sin stocks are less held by norm-constrained institutions such as pension plans as compared to mutual or hedge funds that are natural arbitrageurs, and they receive less coverage from analysts than do stocks of otherwise comparable characteristics. Sin stocks also have higher expected returns than otherwise comparable stocks, consistent with them being neglected by norm-constrained investors and facing greater litigation risk heightened by social norms. Evidence from corporate financing decisions and the performance of sin stocks outside the US also suggest that norms affect stock prices and returns.



## 1. Introduction

Many social scientists believe that social norms are important in shaping economic behavior and market outcomes, overriding at times even the profit motive.[1] An early articulation of this viewpoint in economics is Becker (1957) model of discrimination. In his model, agents (e.g., employers) with discriminatory tastes arising from community norms pay for those tastes by bearing financial costs from their decisions to not interact with particular types of people. Arrow (1972) points out that a complete theory of discrimination also must explain why entrepreneurs without discriminatory tastes cannot make profits by hiring labor cheaply from the groups discriminated against by other employers. Subsequent theories of social norms (e.g., Akerlof, 1980; Romer, 1984) provide sufficient conditions under which social customs that are disadvantageous to the individual nevertheless may persist if individuals are sanctioned by loss of reputation for disobedience of the custom.[2] Empirical work on the

☆ We thank Malcolm Baker (referee) and an anonymous referee for many helpful comments. We also thank Antti Petajisto, Murray Carlson, Douglas Diamond, Kenneth French, Lorenzo Garlappi, Rob Heinkel, Narasimhan Jegadeesh, Lisa Kramer, Alan Kraus, Arvind Krishnamurthy, Jeffrey Kubik, Owen Lamont, Kai Li, Andrew Metrick, Jose Scheinkman, Anna Scherbina, Jeremy Stein, Andrei Ukhov, Rossen Valkanov, Sunil Wahal, Jialin Yu, and seminar participants at the AFA, Emory, McGill, Rutgers, Simon Fraser, Society of Quantitative Analysts, Swedish Institute for Financial Research, the European Finance Association Conference, the Financial Economics and Accounting Annual Conference, the Maryland Behavioral Finance Symposium, the NBER Behavioral Finance Conference, the Pacific Northwest Finance Conference, and the UBC Summer Conference for a number of helpful comments. Kacperczyk acknowledges research support from the Social Sciences and Humanities Research Council of Canada. Address inquiries to hhong@princeton.edu and mkacperc@stern.nyu.edu.
* Corresponding author.
*E-mail addresses:* hhong@princeton.edu (H. Hong), mkacperc@stern.nyu.edu (M. Kacperczyk).

0304-405X/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jfineco.2008.09.001

[1] There are various definitions of the concept of a social norm or custom. Following Akerlof (1980), we define 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.

[2] See Elster (1989) for a review of social norms and economic theory.

effects of social norms on markets has traditionally focused upon measuring the extent of discrimination in the labor market.[3] A related literature points out that social interactions (or peer effects) more generally are important for a variety of economic outcomes (see, e.g., Glaeser and Scheinkman, 2003).[4]

In this paper, we provide new evidence on the market effects of social norms in the novel setting of the stock market. Specifically, we study the investing environment of "sin" stocks, i.e., publicly traded companies involved in the production of alcohol, tobacco, and gaming. This is an ideal setting in which to study the effects of social norms on markets for several reasons. First, there is clearly a societal norm against funding operations that promote human vice, and consequently many investors may not want themselves or others to support these companies by investing in their stocks. Anecdotal evidence supporting this premise can be found in the embrace of socially responsible investing (SRI) by managers of institutions such as pension funds and endowments who screen their investments to rule out sinful stocks such as alcohol, tobacco, and gaming companies. The Social Investment Forum estimates that about $2.34 trillion dollars in 2001 or roughly 12% of the total assets under management in that year undergo some kind of social screens (see Geczy, Stambaugh, and Levin, 2003), which suggests a potentially sizeable effect of socially responsible investing on the prices of sin stocks. This figure has remained fairly constant ever since the Social Investment Forum started making these estimates in the mid-1990s. Second, the stock market provides us with a rich set of data on investor behavior, stock pricing, and firm behavior, which allows us to discriminate more finely among alternative hypotheses than do existing empirical studies of social norms.

A third reason why the stock market is ideally suited for an investigation of market effects from social norms is that there can be significant financial costs associated with norm-constrained investing, i.e., investors pay for their discriminatory tastes à la Becker. To begin with, there is the cost of being unable to diversify into publicly traded sin companies, although this cost is small since there are few of these firms relative to the universe of stocks. More importantly, as we show below, sin stocks tend to be relatively cheap (i.e., with low price-to-book or price-to-earnings ratios) when benchmarked against comparables. Finally, Geczy, Stambaugh, and Levin (2003) find that from the perspective of an investor who seeks to create an optimal portfolio from mutual funds, limiting oneself to funds that include social objectives in their investment policies can be very costly. For instance, an investor who believes that returns are generated by a multifactor pricing model can incur a certainty-equivalent cost of 30 basis points per month, while an investor who believes in managerial skill can incur a cost of more than 100 basis points per month. Importantly, these calculations take as given that there are few publicly traded sin stocks in the marketplace. In reality, the number of sin stocks is likely to be endogenous, depending on the degree to which investors shun them because of social norms.

We begin our investigation of social norm effects on investments in sin stocks by looking at who owns these stocks. First, we hypothesize that the shares of sin stocks should be held in smaller proportions by institutions subject to social norm pressures. These include institutions whose positions in stocks are public information, institutions with diverse constituents, and institutions that can be readily exposed to public scrutiny (e.g., picketing by an unhappy minority). Examples include pension funds, universities, religious organizations, banks, and insurance companies. This hypothesis further implies that sin stocks should be less followed by sell-side analysts who produce financial reports and analyses on companies, since these analysts tend to cater to institutional investors. In contrast to institutional investors, individual investors can keep their stock positions out of the view of enforcers of societal norms, and therefore, we expect individual investors to be more willing than institutional investors to hold sin stocks. Mutual funds and hedge funds represent another class of investors whom we expect to be willing to invest in sin stocks, since they are natural arbitrageurs in the marketplace. While even mutual funds and hedge funds may be increasingly subject to social norm pressures as witnessed by the recent growth of the socially responsible investment class, we expect some of them to flout social conventions and buy sin stocks if those stocks are neglected by other investors and priced cheaply.[5]

Consistent with these predictions, we find that sin stocks have less institutional ownership, as compared to stocks of otherwise comparable characteristics during the period of 1980–2006 for which data are available. Our identification strategy throughout comes from judging sin stock outcomes (e.g., institutional ownership, stock returns, etc.) relative to carefully chosen industry comparables and controls of stock characteristics, i.e., our effects are coming from sin status as opposed to unobserved heterogeneity related to industry or other stock characteristics. Using our most conservative estimates, the sin stock comparables, defined as those with similar Fama and French (1997) industry groupings as our sin stocks have on average about 28% of their shares held by institutions. In contrast, sin stocks have about 23% of their shares held by institutions, which is approximately an 18% lower institutional ownership ratio than that of their comparables.

[3] For a survey of this literature, see Altonji and Blank (1999); and for a recent summary of this work, see Levitt (2004). Many papers have developed clever empirical approaches to identifying the extent of racial discrimination in the labor market. However, the evidence to date supporting taste-based discrimination (as opposed to rational information-based theories) has been mixed.

[4] A growing body of empirical research speaks to the importance of peer effects in a variety of contexts (e.g., Case and Katz, 1991; Glaeser, Sacerdote, and Scheinkman, 1996). Recent work on social interaction and financial markets focuses primarily on investor behavior (see, e.g., Hong, Kubik, and Stein, 2004). There are few papers on whether peer effects have price implications.

[5] An example of such a fund is the VICE mutual fund, which openly promotes holding vice stocks; however, these examples are rare and typically small in size.

In addition, sin stocks receive less analyst coverage during the period of 1976–2006 for which data are available. The typical sin stock comparable in our sample receives coverage from about 1.7 analysts. Sin stocks on average are followed by 1.3 analysts, representing a 21% decline in coverage relative to the mean. These figures have remained relatively stable throughout our sample.

When we break down this analysis by types of institutions (banks, insurance companies, mutual funds, independent investment advisors, and others, such as pension plans and universities), we find that shares of sin stocks are not held in smaller proportions by mutual funds and independent investment advisors, who are the natural arbitrageurs among these institutions.[6] This set of disaggregated results suggests that the low institutional ownership of sin stocks cannot be explained by a story in which institutions are smarter than individuals, since the result does not hold for the mutual funds and hedge funds that are likely to be the smartest investors among these institutions.

Next, we formulate our predictions about the effects of social norms on the returns to investing in sin stocks. From the work of Merton (1987) on neglected stocks and segmented markets, there are at least two reasons why sin stocks should be cheaper than other stocks and hence outperform comparables, even after accounting for well-known predictors of stock returns. First, the neglect of sin stocks by an important set of investors, such as institutions, means that the prices of those stocks will be depressed relative to their fundamental values because of limited risk sharing and hence, sin stocks should have higher expected returns than comparables. Second, because of neglect or limited risk sharing, Merton shows that the CAPM no longer holds and idiosyncratic risk and not just beta matters for pricing. As a result, the increased litigation risk associated with the products of sin companies, which is further heightened by social norms, should further increase the expected returns of sin stocks. For example, tobacco companies faced substantial litigation risk until their settlement with state governments in 1997. Moreover, many practitioners believe that investors may simply underestimate the value of sin stocks. For instance, Berman (2002) writes that "…sin stocks come with other advantages besides stability. Most of these stocks have lower valuations than the overall market. Some of them also offer excellent dividends. …Finally, sin stocks tend to benefit from very conservative accounting because their industries fall under considerable scrutiny from regulators".[7]

Implicit in the neglect-effect hypothesis are two assumptions. The first is limited arbitrage, i.e., not enough arbitrage capital is brought to bear on sin stocks because of a set of constraints and risks articulated by Shleifer and Vishny (1997) and others.[8] The second is that the neglect of sin stocks by institutional investors has been relatively stable over our sample.[9] These assumptions seem reasonable in light of our evidence on institutional holdings.

We test our predictions of sin stocks outperforming comparables by analyzing both prices and returns. First and most conventionally, using time-series regressions during the period of 1965–2006, we find that a portfolio long sin stocks and short their comparables has a return of 26 basis points per month after adjusting for a four-factor model comprising the three Fama-French factors and the momentum (returns) factor. Second, using cross-sectional regressions controlling for firm characteristics (data only available from 1965 to 2006), we find that sin stocks outperform their comparables by 29 basis points a month—again a statistically and economically sizeable magnitude—even after accounting for well-known determinants of expected returns in cross-sectional regressions such as market size, past return, and market-to-book ratio. Our results are robust to the exclusion of tobacco stocks, which arguably might be driving some of these high returns due to both litigation risk and perhaps unexpectedly positive results from litigation. The sin stocks, net of tobacco, have a significant 21 basis points a month of outperformance or 2.5% per year.[10]

Third, we compare the valuation ratios (e.g., market-to-book) of sin stocks to those of other stocks. The valuation ratios of sin stocks are on average about 15–20% lower than those of other companies after controlling for differences in other stock characteristics, an economically and statistically significant effect during the period of 1965–2006. These valuation ratios, using a Gordon growth model calibration, imply excess returns of about 2% a year. These implied annual return numbers are not statistically different from the return figures obtained from our most conservative cross-sectional regressions. Our analysis of sin stock prices using price ratios is more conservative than our analysis of returns, since the realized returns of certain sin stocks (tobacco) may be influenced by unexpectedly good cash flow news over our sample period.

We then consider a number of robustness checks. Most importantly, we extend our analysis of sin stocks outside of the US to the seven large markets in Europe and to Canada. These countries have similar attitudes towards sin stocks. We find that sin stocks in these markets do outperform other stocks by about 2.5% a year, very similar to our conservative estimates in the US and the result is significant at the 10% level of significance. This provides strong out-of-sample support for our hypothesis.

We also relate our empirical findings to calibrations done by Petajisto (2009) and a model adapted from

[6] The category of independent investment advisor is a hodge-podge of different institutions that includes hedge funds. Accordingly, it makes sense for us to include them with mutual funds as natural arbitrageurs.

[7] See 〈http://www.moneysense.ca/shared/print.jsp?content=20021127_154845_3424〉 to retrieve the article.

[8] In other words, Arrow (1972) is right to point out the opportunity for entrepreneurial investors to exploit the discriminatory tastes of other investors, but there are not enough entrepreneurial investors to completely eliminate the impact of discriminatory tastes on stock prices.

[9] We thank the referees for pointing this out to us.

[10] One caveat is that even this might not be conservative enough since one could also argue that the gaming industry might have had unexpectedly good news during our sample period.

Heinkel, Kraus, and Zechner (2001), henceforth HKZ, who develop a model to consider the price implications of ethical investing that excludes companies that pollute. In our context, it is companies that produce sin. Petajisto and HKZ develop models in the spirit of Merton that look at the price implications of limited risk sharing due to neglect induced by social norms or ethical investing. Their price implications are precisely the ones tested here. Our empirical findings match well with their calibration results, which we describe in more detail below.

We further validate our characterization of the prices of sin stocks as being influenced by social norms by looking at the corporate decisions of sin companies. And as a consequence of stock underpricing due to social norms, we predict that sin companies should finance their operations using relatively more debt than equity, since debt markets tend to be less transparent than equity markets as we argue below. Using data from 1962 to 2006, we confirm that sin companies have significantly higher leverage after accounting for the usual predictors of capital structure.

Our findings on the effects of social norms in the context of the stock market strongly support the viewpoint that social norms can have important consequences for markets. Indeed, our results most likely represent lower bounds on the effects of social norms in the stock market, since many companies operating in sin industries may not become public precisely because they are shunned by many investors.

Our paper is related to the work of Teoh, Welch, and Wazzan (1999), who examine the effect of the shareholder boycott of South Africa's apartheid regime. They find that for all the visibility associated with the boycott, there was little discernible effect either on the valuations of banks and corporations with South African operations or on the South African financial markets, because corporate involvement in South Africa was small in the first place. However, they do find some weak evidence that institutional shareholdings in corporations with South African investments increased when those corporations divested.

The rest of our paper proceeds as follows. In Section 2, we provide a background history of sin stocks and describe how we identify these stocks in our analysis. In Section 3, we complete the description of our data set and specify the variables for our regression models. In Section 4, we report the empirical results from our regression analyses. We conclude in Section 5. And details on our calibration exercises are in Appendix A.

## 2. Background on and selection of sin stocks

Our analysis of social norm effects on the stock market focuses on the industries collectively known as the "Triumvirate of Sin", namely alcohol, tobacco, and gaming. Today, all three of these consumer product groups are viewed as sinful by many individuals and social groups in the United States and other countries, due to their addictive properties and undesirable social consequences when consumed excessively. While the sinful aspects of alcohol and gaming have long been recognized by societies into which they have been introduced, tobacco has been the subject of negative social norms only as recently as the past four decades. Our analysis will exploit the time variation in the social norms governing tobacco consumption to test various predictions of social norm effects in financial markets.

Tobacco consumption has been viewed as sinful for only a relatively short period since its introduction to Europe in the mid-16th century.[11] This is because the adverse individual and public health consequences of smoking tobacco were not widely known until the mid-1960s. Indeed, most early European physicians subscribed to the Native American belief that tobacco might be an effective medicine. As cigarette smoking grew in prevalence during the early 20th century, however, articles addressing the adverse health effects of smoking began to appear in scientific and medical journals. Studies beginning in 1930 led the American Cancer Society to warn by 1947 (when other reports began to surface) about the ill effects of smoking, though most of the public remained unaware of the statistical correlation between smoking and cancer. This changed, however, in 1952 when *Reader's Digest* published an article entitled "Cancer by the Carton", which detailed the dangers of smoking.

Although the tobacco industry responded with marketing and studies of their own, thus managing to stave off negative public opinion for another decade, this battle came to an end in the early 1960s with the formation of the Surgeon General's Advisory Committee on Smoking and Health. Convened in response to political pressures and a growing body of scientific evidence, the committee released a 387-page report in 1964 unequivocally concluding that cigarette smoking is causally related to lung cancer. In 1965, Congress passed the Federal Cigarette Labeling and Advertising Act requiring surgeon general's warnings on all packages, and in 1971, all broadcast advertising of tobacco products was banned.

Like alcohol, gambling has long been considered a vice and a sinful activity that corrupts society. In most societies, gambling is heavily regulated because of concerns about criminal involvement. Since the mid-to-late 1990s, however, a number of states in the United States have deregulated casino gaming to legalize its production outside of Native American casinos, under pressure from both political referendums and budget concerns. According to a report from the National Gambling Commission, by 1999, 26 states had joined Nevada and Atlantic City (New Jersey) in legalizing casino-style gaming.[12] Hence, one could argue that gaming has become more socially acceptable in recent years, though many surveys of individuals across various states indicate that public opinion still regards gambling as sinful behavior.[13]

Based upon the Triumvirate of Sin described above, we identify sin stocks from the universe of stocks in the

[11] The following material on tobacco is drawn from a CNN web site profiling the tobacco industry, ⟨http://edition.cnn.com/US/9705/tobacco/history⟩.

[12] For more information regarding this report and discussion of legislative changes in gaming, see Chen and Bin (2001).

[13] See, for instance, surveys reported in a Congressional report on gaming that may be found online at ⟨http://www.library.ca.gov/CRB/97/03/⟩.

following manner. We start with the Fama and French (1997) classification of stocks based upon their SIC codes into 48 industries. Stocks in Fama-French industry group 4 (beer or alcohol) and industry group 5 (smoke or tobacco) are classified as sin stocks. Stocks with SIC codes 2100–2199 belong to the beer group, and those with SIC codes of 2080–2085 are in the smoke group. Unfortunately, the Fama-French classification scheme does not separate gaming stocks from hotel stocks or other entertainment stocks. To this end, we need to use the NAICS classification, which identifies gaming stocks as those bearing the following NAICS codes: 7132, 71312, 713210, 71329, 713290, 72112, and 721120. In sum, sin stocks in our analysis comprise the union of the Fama and French (1997) industry groups 4 and 5 along with the NAICS group for gaming. We work with this expanded 49-industry group throughout our analysis.

We then augment this list by searching across companies at the segment level, as follows.

We utilize the Compustat Segments data, available from 1985 to 2006, which contain information on the SIC and NAICS codes of the different segments of a company. We identify a company as a sin stock if any of its segments has an SIC code in either the beer or the smoke group or an NAICS code in the gaming group, as defined above. Accordingly, our final list of sin stocks is the union of two screening procedures—one applying the Fama and French (1997) and NAICS classifications at the company level and the second applying the same classifications at the segment level using Compustat Segments data. The latter screen on company segment information is essential for obtaining an accurate list of sin stocks, since many companies such as Philip Morris, now known as Altria, have diversified operations. We are unable to implement the augmented search for stocks no longer in existence by 1985, since the segments data are available only after 1985. For stocks listed before 1985 and still in existence after 1985, we back-fill this augmentation procedure—i.e., a stock identified as sinful using the segments data will be characterized as sinful throughout its history.

A list of the sin stocks identified through the above screening procedure, containing Center for Research in Security Prices (CRSP) PERMNO, company name, and the time period of coverage in our data set, can be obtained from the authors' Web sites. The list contains many well-known names such as Altria, Anheuser Busch Co., Bacardi, Bally, Caesars, Loews, Mandalay, and Trump Hotels. We also performed some cross-checks of this list with searches of popular books on sin investing (e.g., Ahrens, 2004; Waxler, 2004) and various online sources, to confirm the accuracy of this list. By and large, our screening procedure appears to yield an accurate and comprehensive list of stocks in the alcohol, tobacco, and gaming industries.

Table 1 provides a more systematic year-by-year look at our data set of sin stocks beginning in 1926 and ending in 2006.[14] There are a total of 193 distinct names, comprised of 36 distinct tobacco companies, 62 distinct

[14] Importantly, in our time-series tests that span a longer horizon, tobacco stocks are considered sinful only in the post-1965 period.

**Table 1**
Profile of sin stocks.

This table reports summary statistics about the sin stocks. In Panel A, we report year-by-year the number of sin stocks that fall into the three subgroups of tobacco, alcohol, and gaming. In Panel B, we report market betas calculated using value-weighted monthly returns on 48 Fama and French (1997) industry portfolios along with our 49th gaming portfolio. The data span the period of 1926–2006.

Panel A: Distribution by year

| Year | All | Tobacco | Alcohol | Gaming |
|---|---|---|---|---|
| 1926 | 18 | 15 | 3 | 0 |
| 1927 | 21 | 18 | 3 | 0 |
| 1928 | 20 | 17 | 3 | 0 |
| 1929 | 21 | 17 | 4 | 0 |
| 1930 | 21 | 17 | 4 | 0 |
| 1931 | 21 | 17 | 4 | 0 |
| 1932 | 21 | 17 | 4 | 0 |
| 1933 | 22 | 16 | 6 | 0 |
| 1934 | 22 | 16 | 6 | 0 |
| 1935 | 23 | 16 | 7 | 0 |
| 1936 | 25 | 16 | 9 | 0 |
| 1937 | 25 | 16 | 9 | 0 |
| 1938 | 25 | 16 | 9 | 0 |
| 1939 | 25 | 16 | 9 | 0 |
| 1940 | 25 | 16 | 9 | 0 |
| 1941 | 23 | 14 | 9 | 0 |
| 1942 | 23 | 14 | 9 | 0 |
| 1943 | 23 | 14 | 9 | 0 |
| 1944 | 23 | 14 | 9 | 0 |
| 1945 | 24 | 14 | 10 | 0 |
| 1946 | 26 | 14 | 12 | 0 |
| 1947 | 27 | 14 | 12 | 1 |
| 1948 | 27 | 13 | 13 | 1 |
| 1949 | 26 | 12 | 13 | 1 |
| 1950 | 25 | 11 | 13 | 1 |
| 1951 | 25 | 11 | 13 | 1 |
| 1952 | 24 | 11 | 12 | 1 |
| 1953 | 25 | 11 | 13 | 1 |
| 1954 | 25 | 11 | 13 | 1 |
| 1955 | 25 | 11 | 13 | 1 |
| 1956 | 25 | 11 | 13 | 1 |
| 1957 | 24 | 10 | 13 | 1 |
| 1958 | 23 | 9 | 13 | 1 |
| 1959 | 23 | 10 | 12 | 1 |
| 1960 | 23 | 10 | 12 | 1 |
| 1961 | 23 | 10 | 12 | 1 |
| 1962 | 23 | 10 | 12 | 1 |
| 1963 | 25 | 10 | 13 | 2 |
| 1964 | 26 | 11 | 13 | 2 |
| 1965 | 26 | 11 | 13 | 2 |
| 1966 | 26 | 11 | 13 | 2 |
| 1967 | 27 | 11 | 14 | 2 |
| 1968 | 27 | 11 | 14 | 2 |
| 1969 | 30 | 11 | 16 | 3 |
| 1970 | 32 | 12 | 16 | 4 |
| 1971 | 32 | 12 | 16 | 4 |
| 1972 | 38 | 12 | 20 | 6 |
| 1973 | 41 | 12 | 22 | 7 |
| 1974 | 41 | 12 | 22 | 7 |
| 1975 | 42 | 12 | 23 | 7 |
| 1976 | 42 | 12 | 23 | 7 |
| 1977 | 42 | 12 | 23 | 7 |
| 1978 | 45 | 12 | 23 | 10 |
| 1979 | 48 | 12 | 23 | 13 |
| 1980 | 50 | 12 | 24 | 14 |
| 1981 | 55 | 12 | 24 | 19 |
| 1982 | 56 | 12 | 25 | 19 |
| 1983 | 58 | 11 | 26 | 21 |
| 1984 | 62 | 12 | 27 | 23 |
| 1985 | 62 | 11 | 27 | 24 |
| 1986 | 60 | 9 | 26 | 25 |
| 1987 | 59 | 9 | 24 | 26 |

Table 1 (*continued*)

Panel A: Distribution by year

| Year | All | Tobacco | Alcohol | Gaming |
|---|---|---|---|---|
| 1988 | 60 | 9 | 24 | 27 |
| 1989 | 61 | 9 | 25 | 27 |
| 1990 | 64 | 9 | 23 | 32 |
| 1991 | 67 | 10 | 23 | 34 |
| 1992 | 70 | 11 | 21 | 38 |
| 1993 | 90 | 11 | 23 | 56 |
| 1994 | 98 | 11 | 22 | 65 |
| 1995 | 106 | 13 | 25 | 68 |
| 1996 | 110 | 15 | 30 | 65 |
| 1997 | 116 | 16 | 32 | 68 |
| 1998 | 107 | 13 | 31 | 63 |
| 1999 | 98 | 13 | 32 | 53 |
| 2000 | 82 | 11 | 29 | 42 |
| 2001 | 73 | 9 | 26 | 38 |
| 2002 | 69 | 7 | 25 | 37 |
| 2003 | 65 | 7 | 23 | 35 |
| 2004 | 63 | 8 | 21 | 34 |
| 2005 | 56 | 8 | 17 | 31 |
| 2006 | 56 | 8 | 17 | 31 |
| Total | 193 | 36 | 62 | 95 |

Panel B: Industry market betas: 1926–2006

| Industry | Beta | Industry | Beta |
|---|---|---|---|
| Agriculture | 0.92 | Guns | 0.84 |
| Food | 0.74 | Gold | 0.65 |
| Soda | 0.82 | Mines | 0.93 |
| **Beer** | **0.94** | Coal | 0.79 |
| **Smoke** | **0.63** | Oil | 0.86 |
| Toys | 1.22 | Utilities | 0.80 |
| Fun | 1.39 | Telecommunication | 0.64 |
| Books | 1.07 | Personal services | 1.07 |
| Household | 0.93 | Business services | 0.94 |
| Clothes | 0.78 | Computers | 1.10 |
| Healthcare | 1.18 | Chips | 1.37 |
| Medical equipment | 0.84 | Laboratory equipment | 0.99 |
| Drugs | 0.85 | Paper | 1.42 |
| Chemicals | 1.02 | Boxes | 0.94 |
| Rubber | 1.07 | Transport | 1.14 |
| Textiles | 1.15 | Wholesalers | 1.10 |
| Building materials | 1.13 | Retailers | 0.96 |
| Construction | 1.35 | Meals | 0.97 |
| Steel | 1.33 | Banks | 1.02 |
| Fabricated products | 1.08 | Insurance | 1.11 |
| Machinery | 1.23 | Real estate | 1.24 |
| Electrical equipment | 1.28 | Financials | 1.26 |
| Autos | 1.22 | Other | 1.05 |
| Aero | 1.32 | **Gaming** | **1.12** |
| Ships | 1.13 | | |

alcohol companies, and 95 distinct gaming companies. There are not a lot of sin stocks—only 56 even in 2006—relative to the thousands of stocks in the universe. While the numbers of tobacco and alcohol stocks have stayed relatively constant through the years, the number of gaming stocks increased significantly in the 1990s with the ongoing deregulation of the gaming industry. As we mentioned earlier, this recent trend suggests that gaming may be gaining more widespread social acceptance. As we also indicated, however, there are a number of other factors underlying this trend.

We further characterize these sin stocks in Panel B of Table 1 by calculating the market betas for the period of 1926–2006 of the different industry portfolios, using data from Fama and French (1997). The market betas are calculated using the time-series of monthly returns on the 49 (value-weighted) industry portfolios (the original 48 Fama and French, 1997, industries plus gaming).[15] The three portfolios of interest (beer, smoke, and gaming) have betas of 0.94, 0.63, and 1.12, respectively. As suggested by practitioners, beer and smoke appear to have somewhat lower betas than other industries during the period of 1926–2006 whereas gaming stocks have betas comparable to those of many other industries.

While we focus on the so-called Triumvirate of Sin (alcohol, tobacco, and gaming), two other classes of stocks are sometimes thought of as sinful. The first is the sex industry. However, there are very few publicly traded companies with heavy operations in sex.[16] Therefore, omitting these companies will not affect our results in any significant way. The second is the defense industry. We have decided against including defense as a sin industry in our main analysis because it is not clear that defense is considered a sin by many Americans. As a robustness exercise, however, we later broaden our definition of sin stocks to include stocks in Fama and French (1997) industry grouping 26 (guns).

## 3. Data

Having described our procedure for identifying sin stocks, we now characterize our data collection on the universe of stocks. Our data on US firms come from CRSP and Compustat. From CRSP, we obtain daily closing stock prices, daily shares outstanding, and daily dollar trading volumes for NYSE, Amex, and Nasdaq stocks over the period of 1962–2006. From Compustat, we obtain annual information on a variety of accounting variables during the same period. To be included in our sample, a firm must have the requisite financial data from both CRSP and Compustat. We follow other studies in focusing on companies with CRSP share codes of 10 or 11 and excluding firms with one-digit SIC codes of 6, which belong to the financial services industry.[17]

Our data on ownership structures come from the CDA Spectrum Database of 13-F filings by institutional investors, defined as those managing at least $100 million in assets. This database reports holdings of a particular stock in terms of shares held by various classes of institutional

[15] Note that in calculating these betas, we have not excluded the gaming stocks from the Fama and French portfolios. This is unlikely to bias the estimates of beta for the different industries by a significant amount, since there are so few gaming stocks.

[16] Two public companies with large sex operations, according to practitioners, are Playboy Enterprises and the Barcelona-based Private Media Group. However, many other large entertainment conglomerates profit off sex on cable television through holdings in various subsidiaries that are difficult to track down.

[17] Another screen sometimes used in empirical analyses over the universe of stocks is to drop firms with book values of <$10 million. We have replicated our analyses below using this additional screen and found similar results.

investors. The five institution types are banks, insurance companies, mutual funds, independent investment advisors (which includes hedge funds), and others (including universities, pension plans, and employee ownership plans). While the 13-F filing is available quarterly, most companies only file timely reports on a semi-annual basis, at the end of June and at the end of December. Our analysis will focus on the end-of-year filings. Our data on analyst coverage come from the Institutional Brokers Estimates System (IBES) database, which reports the number of analyst estimates of earnings issued on a stock at various points in time (typically quarterly).

### 3.1. Variables in ownership regressions

Institutional ownership ($IO_{it}$) is the fraction of the shares of company $i$ held by institutions in the CDA Spectrum Database at the end of year $t$. $IO$ is calculated by aggregating the shares held by all five types of institutions at the end of the year and then dividing this amount by shares outstanding at the end of the year (Item 25 in Compustat). $LOGSIZE_{it}$ is the natural logarithm of firm $i$'s market capitalization (price times shares outstanding) at the end of year $t$. $LOGMB_{it}$ is the natural logarithm of firm $i$'s market cap divided by its book value at the end of year $t$. $STD_{it}$ is the standard deviation of daily (simple, raw) returns during year $t$. $BETA_{it}$ is the beta of firm $i$'s industry, among the 49 industries listed in Table 1, in year $t$. $PRINV_{it}$ is the inverse of firm $i$'s share price at the end of year $t$. $RET_{it}$ is the average monthly return on stock $i$ during year $t$.

The summary statistics for these variables are provided in Panel A of Table 2. The time-series average of the cross-sectional means of $IO$ is 0.28, and the time-series average of the cross-sectional standard deviations of $IO$ is 0.29. In other words, in a typical year, a typical firm has about 28% of its shares held by institutions, and the standard deviation of institutional ownership in a typical cross-section is 29%. The other variables are standard and do not merit discussion, except to report that their summary statistics are similar to those found in earlier studies.

### 3.2. Variables in analyst coverage regressions

Our measure of analyst coverage is $LOGCOV_{it}$, defined as the natural logarithm of one plus the number of analysts covering firm $i$ at the end of year $t$. As in earlier studies, stocks that do not appear in IBES are assumed to have no analyst estimates.[18] The other variables used in these regressions are constructed in the same way as those in the ownership regressions. The summary statistics for these variables over the period of 1976–2006 are reported in Panel B of Table 2. Note that the time-series average of

[18] If a stock does not have an analyst estimate at the end of the year, we look backwards through the year to find the most recent estimate and use this as the measure of analyst coverage for the firm during the year. Alternatively, we also have experimented with using October 1 as a cutoff date. In other words, if a stock does not have any analyst estimates in the last quarter, then it is recorded as having zero analyst estimates for the year.

**Table 2**
Summary statistics.

This table reports summary statistics for the variables used for the six sets of regressions. Panel A reports the summary statistics (time-series averages of cross-sectional means and standard deviations) for the institutional ownership regressions. Institutional ownership (*IO*) is the fraction of shares of a firm held by institutions. *LOGSIZE* is the logarithm of the market capitalization of the company. *LOGMB* is the logarithm of the market-to-book variable. *STD* is the daily stock return standard deviation during the past year. *BETA* is the firm's industry market beta. *PRINV* is the inverse of the stock price. *RET* is the arithmetic average of the last year's monthly returns. These variables are calculated at the end of the year. Panel B reports the similar summary statistics for the analyst coverage regressions. *LOGCOV* is the log of one plus the number of analyst estimates issued on a company at the end of the year. The other variables in Panel B are constructed in the same way as in Panel A except over a different time period. Panel C reports the summary statistics for the time-series return regressions. *EXCOMP* is the excess monthly return net of the risk-free rate for an equal-weighted portfolio of sin stocks net of comparable stocks. *EXSINP* is the excess monthly return net of the risk-free rate for an equal-weighted portfolio of sin stocks. *MKTPREM* is the excess monthly return of the value-weighted CRSP index. *SMB* is the return of a portfolio long small stocks and short large stocks. *HML* is the return of a portfolio long high book-to-market stocks and short low book-to-market stocks. *MOM* is the return of a portfolio long past 12-month return winners and short past 12-month return losers. Panel D reports the summary statistics for the cross-sectional return regressions. *EXMRET* is the monthly return of a stock net of the risk-free rate. *LOGSIZE*, *LOGMB*, *TURN*, and *RET* are now calculated at the end of each month. *BETA* is just the firm's industry beta, which is now calculated at the end of each month using the past 36-months of data. *TURN*, calculated at the end of month, is the daily share turnover during the past 12-months. *LOGAGE* is the natural logarithm of the firm's age, measured by the number of years available in the CRSP/Compustat data. *BLEV* is book leverage. Panel E reports the summary statistics for the price-to-book regressions. *LOGPE* is the log of the price-to-earnings ratio. *LOGPEBITDA* is the log of price-to-earnings before taxes and after depreciation ratio. *ROE* is return on equity. *RDSALES* is the fraction of research and development expenditures to firm sales. Panel F reports the statistics for the financing decision regressions. *MLEV* is market leverage. *CASH* is cash balances. *DIVPAY* is the dividend payout ratio. *REP* is share repurchases. *PAYOUT* is the overall firm payout ratio. *TOBQ* is Tobin's Q, which is market value of equity plus assets minus book value of equity over assets. *TANG* is firm tangibility defined as net plant, property, and equipment divided by total assets. *PROFIT* is earnings before interest, taxes, and depreciation divided by total assets. *LOGSALES* is the log of net sales.

Panel A: Institutional ownership regressions: 1980–2006

| Variable | Time-series average of means | Time-series average of standard deviations |
|---|---|---|
| *IO* | 0.28 | 0.29 |
| *LOGSIZE* ('000) | 11.80 | 2.28 |
| *BETA* | 1.24 | 0.45 |
| *LOGMB* | 0.74 | 0.89 |
| *PRINV* | 0.32 | 0.80 |
| *STD* (%) | 3.75 | 2.45 |
| *RET* (%) | 1.25 | 5.65 |

Panel B: Analyst coverage regressions: 1976–2006

| Variable | Time-series average of means | Time-series average of standard deviations |
|---|---|---|
| *LOGCOV* | 0.99 | 1.05 |
| *LOGSIZE* ('000) | 11.21 | 2.19 |
| *BETA* | 1.14 | 0.38 |
| *LOGMB* | 0.61 | 0.93 |
| *PRINV* | 0.37 | 0.86 |
| *STD* (%) | 3.91 | 2.56 |
| *RET* (%) | 1.34 | 5.95 |

Table 2 (*continued*)

Panel C: Time-series return regressions: 1965–2006

| Variable | Mean | Standard deviation |
|---|---|---|
| *EXCOMP* (%) | 0.25 | 2.57 |
| *EXSINP* (%) | 0.96 | 5.62 |
| *MKTPREM* (%) | 0.46 | 4.45 |
| *SMB* (%) | 0.28 | 3.27 |
| *HML* (%) | 0.45 | 2.94 |
| *MOM* (%) | 0.83 | 4.07 |

Panel D: Cross-sectional return regressions: 1965–2006

| Variable | Time-series average of means | Time-series average of standard deviations |
|---|---|---|
| *EXMRET* (%) | 0.87 | 18.17 |
| *LOGSIZE* (‘000) | 11.27 | 2.11 |
| *BETA* | 1.13 | 0.37 |
| *TURN* (%) | 0.36 | 0.58 |
| *LOGMB* | 0.55 | 0.90 |
| *RET* (%) | 1.35 | 4.82 |
| *LOGAGE* | 2.23 | 0.99 |
| *BLEV* (%) | 30.92 | 24.08 |

Panel E: Valuation regressions: 1965–2006

| Variable | Time-series average of means | Time-series average of standard deviations |
|---|---|---|
| *LOGMB* | 0.41 | 0.84 |
| *LOGPE* | 2.73 | 0.98 |
| *LOGPEBITDA* | 2.11 | 1.02 |
| *ROE* (%) | 4.21 | 35.24 |
| *RDSALES* (%) | 10.55 | 92.98 |

Panel F: Corporate financing decisions regressions: 1965–2006

| Variable | Time-series average of means | Time-series average of standard deviations |
|---|---|---|
| *BLEV* | 0.34 | 0.31 |
| *MLEV* | 0.27 | 0.25 |
| *CASH* | 0.17 | 0.28 |
| *PAYOUT* | 0.28 | 1.28 |
| *DIVPAY* | 0.16 | 0.45 |
| *REP* | 0.12 | 0.91 |
| *TOBQ* | 1.92 | 2.08 |
| *TANG* | 0.32 | 0.24 |
| *PROFIT* | 0.07 | 0.26 |
| *LOGSALES* | −0.06 | 0.94 |

the cross-sectional means of *LOGCOV* is 0.99, and the time-series average of the cross-sectional standard deviations is 1.05. Thus, in a typical year, a typical firm has about 1.7 analysts with a standard deviation of about 1.9 analysts. The summary statistics for the other variables in Panel B differ slightly from those in Panel A because of their different sample periods.

### 3.3. Variables in time-series return regressions

In our time-series return regressions, the dependent variable *EXCOMP* ($SINP_t$–$COMP_t$) is the monthly return of an equal-weighted portfolio of sin stocks in month $t$, net of the monthly return of an equal-weighted portfolio of comparable stocks that belong to the Fama and French (1997) industry groups 2 (food), 3 (soda), 7 (fun), and 43 (meals and hotels). *EXSINP* is the monthly return of the sin portfolio net of the risk-free rate. $MKTPREM_t$ is the monthly return of the CRSP value-weighted portfolio in month $t$, net of the risk-free rate. $SMB_t$, $HML_t$, and $MOM_t$ are well-known portfolio return series downloaded from Ken French’s Web site. *SMB* is the monthly return of a portfolio that is long on small stocks and short on large stocks. *HML* is the monthly return of a portfolio that is long on high book-to-market stocks and short on low book-to-market stocks. *MOM* is the monthly return of a portfolio that is long on past one-year return winners and short on past one-year return losers. The summary statistics for the various portfolio returns are presented in Panel C of Table 2.

### 3.4. Variables in cross-sectional return regressions

Our empirical analyses of sin stock returns employ monthly measures of returns as dependent variables. In our cross-sectional return regressions, the dependent variable $EXMRET_{it}$ is the monthly return of an individual stock $i$ in month $t$ net of the risk-free rate. $LOGSIZE_{it}$ is defined as in Panel A, except that we now calculate $LOGSIZE_{it}$ on a monthly as opposed to annual basis. $BETA_{it}$ is the time-varying beta of the industry to which firm $i$ belongs, calculated over the most recent three years of monthly data leading up to and including month $t$. $TURN_{it}$ is the average of daily share turnover in stock $i$—defined as shares traded divided by shares outstanding—during month $t$. $LOGMB_{it}$ is the natural logarithm of the market-to-book ratio of stock $i$ at the end of month $t$. $RET_{it}$ is the average of the most recent 12 months of returns on stock $i$ leading up to and including month $t$. Finally, *LOGAGE* is the natural logarithm of the firm’s age, measured by the number of years available in the CRSP/Compustat data. *BLEV* is the book leverage of the company, which is defined more precisely below. The summary statistics of these variables are reported in Panel D of Table 2.

### 3.5. Variables in valuation regressions

$LOGMB_{it}$ is the natural logarithm of firm $i$’s market-to-book ratio measured at the end of year $t$. $LOGPE_{it}$ is the natural logarithm of the firm’s price-to-earnings ratio at the end of the year. $LOGPEBITDA_{it}$ is the natural logarithm of the firm’s price-to-EBITDA (earnings before interest, tax, depreciation, and amortization). $ROE_{it}$ is firm $i$’s return on equity in year $t$. *ROE* is calculated as the ratio of earnings during year $t$ over the book value of equity at the end of year $t$. Earnings are calculated as income before extraordinary items available to common stockholders (Item 237), plus deferred taxes from the income statement (Item 50), plus investment tax credit (Item 51). $RDSALES_{it}$ is the fraction of firm $i$’s research and development expenditures (Item 46) to firm sales (Item 12) in year $t$. We also introduce $RDMISS_{it}$, which is a dummy variable that equals one if firm $i$’s R&D expenditures observation in year $t$ is missing. $SP500_{it}$ is a dummy variable equal to one if company $i$ is part of the S&P 500 index in year $t$. The time-series average of the cross-sectional means of *LOGMB* is 0.41, and the time-series average of the cross-sectional standard deviations is 0.84. Thus, in a typical year, a typical firm in our sample has a market-to-book ratio of 1.51 with a standard deviation of about 2.32. The summary statistics for these variables are reported in Panel E of Table 2.

### 3.6. Variables in corporate financing decision regressions

The book leverage of firm $i$ in year $t$, denoted by $BLEV_{it}$, is total debt divided by the sum of total debt and book equity ((Item 9+Item 34)/(Item 9+Item 34+Item 216)), measured at fiscal year-end. We also use market leverage, denoted by $MLEV_{it}$, which is the same as *BLEV* except that we replace Item 216 with the firm's market capitalization (the average of firm $i$'s market capitalization over calendar year $t$). $CASH_{it}$ is firm $i$'s cash balances (Item 1) divided by book assets at the start of year $t$ (Item 6). For firm $i$ in year $t$, $PAYOUT_{it}$ is calculated as Compustat Item 115 minus preferred stock reduction plus Item 21, all divided by net income (Item 172), where the preferred stock reduction is the maximum between zero and the difference between the previous year's Item 10 and the current year's Item 10. We then break down this payout variable into its constituent parts. $DIVPAY_{it}$ is firm $i$'s dividend payout ratio in year $t$, namely Item 21 divided by Item 172. $REP_{it}$ is firm $i$'s repurchases divided by net income in year $t$ (Item 115 minus preferred stock reduction, all divided by Item 172).

Tobin's $Q$, $TOBQ_{it}$, is the market value of equity (price times shares outstanding from CRSP) plus assets minus the book value of equity (Item 60+Item 74), all divided by assets, measured at the end of year $t$. Asset tangibility, $TANG_{it}$, is defined as firm $i$'s net plant, property, and equipment (Item 8) divided by total assets at the end of year $t$ (Item 6) and expressed in percentage terms. Profitability, $PROFIT_{it}$, is defined as earnings before interest, taxes, and depreciation (Item 13) divided by total assets at the end of year $t$ and expressed in percentage terms. Finally, $LOGSALES_{it}$ is the natural logarithm of net sales of firm $i$ in year $t$ (Item 12).

## 4. Results

### 4.1. Institutional ownership

In this section, we test whether the shares of sin stocks are less held by institutions that are subject to social norm pressures, while controlling for a host of other firm characteristics such as firm size and stock beta. We hypothesize that institutions such as pension funds, universities, religious organizations, banks, and insurance companies are less willing than other types of investors to hold sin stocks due to the public nature of their investments, their diverse constituencies, and their exposure to public scrutiny (e.g., picketing by an unhappy minority).

Our empirical objective is to determine whether sin stocks have a different institutional investor following than do other stocks. Our task is complicated by the fact that sin stocks will differ from other stocks along a number of other dimensions such as firm size and firm beta. Indeed, earlier research has identified some key firm characteristics that are correlated with institutional ownership.[19] For example, one reason why firm size is correlated with ownership by institutions is that institutions tend to avoid small stocks because of liquidity issues. In this paper, we remain agnostic about why certain firm characteristics tend to be correlated with ownership by institutions. Our strategy is to soak up as much of the cross-sectional variation as possible so that we can better identify the differential effect of sin stocks compared to non-sin stocks.

To this end, we estimate the following regression model:

$$IO_{it} = a_0 + a_1\ SINDUM_{it} + \boldsymbol{a}_2\boldsymbol{X}_{it} + \varepsilon_{it}, \quad i = 1, \ldots, N, \qquad (1)$$

where *SINDUM* equals one if the stock is a sin stock and zero otherwise, $\boldsymbol{X}_{it}$ is a vector of firm characteristics, and $\varepsilon_{it}$ is measurement error. $\boldsymbol{X}_{it}$ includes various permutations of the following variables defined earlier: *LOGSIZE*, *BETA*, *LOGMB*, *PRINV*, *STD*, and *RET*. Also included in the set of control variables are *NASD*, a dummy that equals one if the stock is listed on Nasdaq and zero otherwise; and *SP500*, a dummy variable that equals one if the stock is part of the S&P 500 index and zero otherwise. The coefficient of interest is $a_1$, which measures whether sin stocks have a different ownership structure than do other stocks, controlling for the other firm characteristics contained in $\boldsymbol{X}_{it}$. The null hypothesis is that $a_1$ equals zero, whereas our prediction is that it will be significantly $<0$.

In order to interpret the coefficient on the sin stock dummy variable, *SINDUM*, as a test of social norm effects, we must address the possibility that it might pick up other effects associated with the industries in which our sin stocks reside. In particular, our sin stocks tend to be consumer goods, and it is possible that consumer goods differ from other industries in terms of attracting institutional ownership. To eliminate this potentially confounding interpretation, we create two new control variables. The first and more straightforward one is *ONEDIGDUM*, which equals one if a stock resides in the same one-digit SIC code industry as any of our sin stocks and zero otherwise. With the inclusion of *ONEDIGDUM*, our *SINDUM* effect cannot be interpreted in terms of institutions favoring certain one-digit SIC code industries over others.

A more subtle and potentially more conservative way to address the issue of related industry effects on institutional ownership is to find comparables for our sin stocks. Conveniently, the Fama and French (1997) industry groups 2 (food), 3 (soda), 7 (fun), and 43 (meals and hotels) provide natural comparables to our sin stocks. Indeed, in some other industry classification schemes,

[19] For evidence related to predictors of institutional ownership, see Del Guercio (1996) and Gompers and Metrick (2001).

beer is often lumped together with soda, as is tobacco or smoke with food, and gaming with fun and meals. Accordingly, we create the dummy variable *GDUM*, which equals one if a stock resides in the set of Fama and French (1997) industry classifications 2 (food), 3 (soda), 4 (beer), 5 (smoke), 7 (fun), 43 (meals), and our 49th industry (gaming), and zero otherwise. With *GDUM* as a control variable, our sin stock effect cannot be interpreted in terms of institutions favoring other industries over the consumer industries included in the definition of *GDUM*. This careful matching of sin stocks to industry comparables and other stock characteristics is the core of our identification strategy.

We estimate our regression model using the ultra-conservative method of running a pooled (panel) regression and calculating standard errors by clustering at the industry level (using our 49 industry groupings). This approach addresses the concern that the errors, conditional on the independent variables, are correlated within industry groupings (e.g., Moulton, 1986). One reason why this may occur is that if an institution decides to invest in a stock in a particular industry, it may also invest in comparables in the same industry for reasons of liquidity or diversification. We also have estimated our model using the Fama and MacBeth (1973) methodology with Newey and West (1987) standard errors, and we obtain similar results. We report our findings from the estimation with industry-clustered standard errors because this method makes fewer assumptions about how the errors are correlated over time, and thus, it is likely to give us more conservative standard errors.[20]

Our results from estimating various specifications of Eq. (1) are reported in Panel A of Table 3. In column 1, we report our estimates of a model that includes as independent variables *SINDUM*, *ONEDIGDUM*, *LOGSIZE*, *BETA*, *NASD*, and *SP500*. We find that the coefficient in front of *SINDUM* is −0.0672 and is statistically significant at the 1% level of significance. The coefficient in front of *ONEDIGDUM* is negative but tiny and statistically insignificant. The mean institutional ownership in our sample is 28%. There is no difference in institutional ownership between sin stock comparables in the same one-digit SIC code and the rest of the sample. In contrast, sin stocks on average have about 21% of their shares held by institutions, which represents about a 24% shortfall relative to the mean. Both *LOGSIZE* and *BETA* attract positive and statistically significant coefficients, suggesting that institutions seem to favor big firms and firms from industries with high market betas. Finally, Nasdaq stocks and stocks in the S&P 500 index have less institutional ownership as compared to other stocks.[21]

In column 2, we replace *ONEDIGDUM* with *GDUM*. The estimated coefficient on *GDUM* bears a negative sign (−0.0209), suggesting that institutions may shy away from consumer goods relative to other industries, but it is statistically insignificant. However, including *GDUM* rather than *ONEDIGDUM* reduces the economic effect of our *SINDUM* variable. The coefficient in front of *SINDUM* is now −0.0487, but it remains statistically significant at the 10% level of significance. Despite its reduction in magnitude in this version of our model, the economic effect of sin stocks on institutional ownership is still sizeable. The sin stock comparables, defined as those with similar Fama and French (1997) industry groupings as our sin stocks, have on average about 28% of their shares held by institutions. In contrast, sin stocks have about 23% of their shares held by institutions, which is approximately an 18% lower institutional ownership ratio than their comparables.

In columns 3–6, we successively experiment with different permutations of control variables to illustrate the robustness of our finding. In column 3, we include *LOGMB* as an additional control; in column 4, we substitute *PRINV* for *LOGMB* since these two variables are highly correlated, as both are scaled by price. In column 5, we add in *STD* (standard deviation of stock return) and find that it attracts a negative coefficient and is statistically significant, indicating that institutions tend to favor low variance stocks. Finally in column 6, we add in the average monthly return over the past 12 months (*RET*). *RET* attracts a negative coefficient and is also statistically significant. The key observation to make across all of these alternative specifications is that the coefficient in front of *SINDUM* remains statistically significant throughout at the 10% level of significance (except for column 3), with economic effects that are similar to that reported in column 2. For instance, in column 6, sin stocks are still estimated to have a 16% lower institutional ownership ratio than their comparables.[22]

Next, we re-specify the ownership regression to disaggregate the effects of sin stocks on holdings by different classes of institutions. Note that for these regressions the data run only from 1980 to 1997.[23] As described earlier, our data set disaggregates ownership information into five types of institutions: type 1 representing banks, type 2 insurance companies, type 3 mutual funds, type 4 independent investment advisors, and type 5 all other institutions, including universities, employee stock ownership plans, etc. We expect types 3 and 4 institutions to be less constrained by social norms as compared to the other three types of institutions, since types 3 and 4 are the natural arbitrageurs in the market. Accordingly, we divide the institutions in our data set into two subgroups, placing types 1, 2, and 5 in one group and types 3 and 4 in another group. We then create two new dependent variables for the regression specification reported in column 6 of Panel A, Table 3. The first of

[20] Where appropriate, we also performed Tobit regressions. The results of the tests remain qualitatively similar.

[21] Falkenstein (1996) reports similar findings for a sample of mutual funds.

[22] The cross-sectional relationships between ownership and the various firm characteristics examined in this section are also very stable across years. (A table that reports this cross-sectional regression year by year is available upon request.)

[23] The truncation of our sample is necessitated by vendor–distributor compatibility issues in the data on institution types after 1997. After 1997, many institutions in the data are erroneously labeled as type 5 institutions. For completeness, we also estimated a pooled regression that includes data after 1997 and found similar results.

**Table 3**
Institutional ownership and analyst coverage of sin stocks.

In Panel A, the dependent variable is institutional ownership (*IO*), which is calculated at the end of each year. *SINDUM* equals one if a stock is a sin stock (alcohol, gaming, or tobacco) and zero otherwise. *GDUM* equals one if a stock is a sin stock or comes from the Fama and French (1997) industry groupings: 2 (food), 3 (soda), 7 (fun), and 43 (meals), and zero otherwise. *LOGSIZE* is the logarithm of the market capitalization of the company. *BETA* is the firm's industry rolling market beta calculated over the last 36 months. *LOGMB* is the logarithm of the market-to-book variable. *PRINV* is the inverse of the stock price. *STD* is the daily stock return standard deviation during the past year. *RET* is the average monthly return during the past year. *NASD* equals one if the stock is listed on Nasdaq and zero otherwise. *SP500* equals one if the stock is in the S&P 500 index and zero otherwise. These variables are calculated at the end of the year. Panel A reports the results of a pooled regression with Moultons (1986) standard errors clustered at the 49-industry groupings. The ownership data cover the period 1980–2006. Panel B reports the results, in which the dependent variable is ownership by two subgroups. In the first column, the dependent variable is the fraction of shares held by type 1 (banks), type 2 (insurance companies), and type 5 (others including pension plans, endowments, and employee-ownership plans) institutions. In the second column, the dependent variable is shares held by type 3 (mutual funds) and type 4 (independent investment advisors) institutions. *T*-test of the difference of the coefficients in front of *SINDUM* across these two regressions is provided (Greene, 2000). The data for Panel B run only from 1980–1997. In Panel C, the dependent variable is analyst coverage, calculated as a log of one plus the number of analyst estimates issued on a stock at the end of the year (*LOGCOV*). The analyst data are from the period 1976–2006. Panel D reports the results from specification (6) in Panel A for different subperiods. All regressions include one-digit SIC industry fixed effects. *** 1% significance; ** 5% significance; and * 10% significance.

Panel A: Institutional ownership: 1980–2006

| Variable | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| *SINDUM* | −0.0672*** | −0.0487* | −0.0421 | −0.0489* | −0.0449* | −0.0448* |
| | (0.0095) | (0.0269) | (0.0275) | (0.0271) | (0.0252) | (0.0254) |
| *ONEDIGDUM* | −0.0024 | | | | | |
| | (0.0122) | | | | | |
| *GDUM* | | −0.0209 | −0.0216 | −0.0209 | −0.0253 | −0.0257 |
| | | (0.0279) | (0.0278) | (0.0280) | (0.0263) | (0.0265) |
| *LOGSIZE* | 0.0819*** | 0.0818*** | 0.0950*** | 0.0831*** | 0.0765*** | 0.0781*** |
| | (0.0031) | (0.0031) | (0.0039) | (0.0030) | (0.0026) | (0.0028) |
| *BETA* | 0.1204*** | 0.1194*** | 0.1362*** | 0.1209*** | 0.1311*** | 0.1322*** |
| | (0.0367) | (0.0376) | (0.0431) | (0.0372) | (0.0383) | (0.0387) |
| *LOGMB* | | | −0.0406*** | | | |
| | | | (0.0040) | | | |
| *PRINV* | | | | 0.0063* | 0.0172*** | 0.0148*** |
| | | | | (0.0035) | (0.0022) | (0.0025) |
| *STD* | | | | | −1.4783*** | −1.3974*** |
| | | | | | (0.1339) | (0.1241) |
| *RET* | | | | | | −0.1904*** |
| | | | | | | (0.0261) |
| *NASD* | −0.0111* | −0.0113* | 0.0065 | −0.0113* | −0.0023 | −0.0005 |
| | (0.0066) | (0.0064) | (0.0068) | (0.0064) | (0.0066) | (0.0066) |
| *SP500* | −0.0415*** | −0.0414*** | −0.0929*** | −0.0445*** | −0.0380*** | −0.0422*** |
| | (0.0101) | (0.0102) | (0.0108) | (0.0098) | (0.0095) | (0.0098) |

Panel B: Pooled regressions by different types of institutions: 1980–1997

| Variable | (Type 1+2+5) | (Type 3+4) |
|---|---|---|
| *SINDUM* | −0.0270** | −0.0111 |
| | (0.0124) | (0.0134) |
| *GDUM* | −0.0035 | −0.0177 |
| | (0.0101) | (0.0142) |
| *LOGSIZE* | 0.0270*** | 0.0465*** |
| | (0.0010) | (0.0023) |
| *BETA* | 0.0237 | 0.1213*** |
| | (0.0152) | (0.0270) |
| *PRINV* | 0.0064*** | 0.0026* |
| | (0.0006) | (0.0013) |
| *STD* | −0.4492*** | −0.3164*** |
| | (0.0563) | (0.0907) |
| *RET* | −0.1432*** | 0.0019 |
| | (0.0094) | (0.0186) |
| *NASD* | −0.0202*** | 0.0187*** |
| | (0.0023) | (0.0032) |
| *SP500* | 0.0388*** | −0.0441** |
| | (0.0050) | (0.0066) |
| *t*-Test of difference of coefficients in front of *SINDUM* | −0.0159*** | |
| | (0.0064) | |

Table 3 (*continued*)

Panel C: Analyst coverage: 1976–2006

| Variable | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| *SINDUM* | −0.1650** | −0.1395 | −0.1578 | −0.1425* | −0.1443* | −0.1400* |
| | (0.0777) | (0.0880) | (0.1139) | (0.0835) | (0.0834) | (0.0846) |
| *ONEDIGDUM* | 0.0169 | | | | | |
| | (0.0207) | | | | | |
| *GDUM* | | −0.0187 | −0.0207 | −0.0187 | −0.0171 | −0.0212 |
| | | (0.0485) | (0.0434) | (0.0492) | (0.0499) | (0.0492) |
| *LOGSIZE* | 0.3637*** | 0.3637*** | 0.4060*** | 0.3830*** | 0.3857*** | 0.4037*** |
| | (0.0079) | (0.0078) | (0.0079) | (0.0081) | (0.0084) | (0.0084) |
| *BETA* | 0.1066* | 0.1058* | 0.1584*** | 0.1326** | 0.1276** | 0.1414** |
| | (0.0559) | (0.0537) | (0.0452) | (0.0561) | (0.0578) | (0.0568) |
| *LOGMB* | | | −0.1476*** | | | |
| | | | (0.0094) | | | |
| *PRINV* | | | | 0.0872*** | 0.0824*** | 0.0547*** |
| | | | | (0.0074) | (0.0059) | (0.0047) |
| *STD* | | | | | 0.6350** | 1.6675*** |
| | | | | | (0.2895) | (0.3276) |
| *RET* | | | | | | −2.2039*** |
| | | | | | | (0.0709) |
| *NASD* | −0.0849*** | −0.0841*** | 0.0207 | −0.0839*** | −0.0879*** | −0.0633** |
| | (0.0292) | (0.0295) | (0.0294) | (0.0301) | (0.0298) | (0.0298) |
| *SP500* | 0.3089*** | 0.3097*** | 0.2471*** | 0.2680*** | 0.2654*** | 0.2250*** |
| | (0.0325) | (0.0326) | (0.0310) | (0.0321) | (0.0325) | (0.0326) |

Panel D: Ownership regressions by subperiods

| Variable | 1980–1984 | 1985–1989 | 1990–1994 | 1995–1999 | 2000–2006 |
|---|---|---|---|---|---|
| *SINDUM* | −0.0660*** | −0.0163 | −0.0305* | −0.0513** | −0.0638 |
| | (0.0137) | (0.0175) | (0.0170) | (0.0227) | (0.0606) |
| *GDUM* | 0.0041 | −0.0220 | −0.0375** | −0.0151 | −0.0325 |
| | (0.0175) | (0.0155) | (0.0166) | (0.0247) | (0.0589) |
| *LOGSIZE* | 0.0581*** | 0.0631*** | 0.0708*** | 0.0734*** | 0.0889*** |
| | (0.0039) | (0.0025) | (0.0035) | (0.0037) | (0.0046) |
| *BETA* | 0.0416 | 0.0921*** | 0.0698** | 0.0489 | 0.0311 |
| | (0.0374) | (0.0326) | (0.0337) | (0.0427) | (0.0189) |
| *PRINV* | 0.0136*** | 0.0062*** | 0.0075*** | −0.0033 | −0.0386** |
| | (0.0024) | (0.0016) | (0.0019) | (0.0126) | (0.0150) |
| *STD* | −1.3000*** | −0.8690*** | −0.5989*** | −1.3479*** | −1.8796*** |
| | (0.2547) | (0.1376) | (0.0957) | (0.2004) | (0.3205) |
| *RET* | −0.0799* | −0.2751*** | −0.1176*** | −0.2177*** | −0.2314*** |
| | (0.0428) | (0.0347) | (0.0241) | (0.0235) | (0.0595) |
| *NASD* | −0.0222*** | −0.0351*** | −0.0192*** | −0.0126 | 0.0257*** |
| | (0.0051) | (0.0051) | (0.0065) | (0.0081) | (0.0094) |
| *SP500* | 0.0467*** | 0.0159 | −0.0028 | −0.0058 | −0.1164*** |
| | (0.0128) | (0.0096) | (0.0112) | (0.0135) | (0.0179) |

these is the fraction of shares held by types 1, 2, and 5 institutions, and the second is the fraction of shares of a company held by types 3 and 4 institutions.

Our findings from estimating the column 6 regression specification on the new disaggregated measures of institutional ownership are reported in Panel B of Table 3. Consistent with our earlier findings for all institutions, for the types 1, 2, and 5 institutions the coefficient in front of *SINDUM* is negative (−0.0270) and statistically significant at the 5% level of significance. In contrast, the coefficient in front of *SINDUM* for the types 3 and 4 institutions is smaller in absolute magnitude (−0.0111) and statistically insignificant. Indeed, a *t*-test of the difference in the magnitudes between these two coefficients (−0.0159) is statistically significant at the 1% level of significance. In other words, sin stocks are not less held by mutual funds and independent investment advisors. This fits well with our hypothesis that mutual funds and hedge funds are the most likely types of institutions to play the role of arbitrageurs and buy sin stocks if they are ignored and priced cheaply.

### 4.2. Analyst coverage

In light of our finding that sin stocks have lower institutional ownership as compared to other stocks, logic suggests that sin stocks should also be less followed by sell-side analysts who produce financial reports and analyses on companies. Our take on analyst coverage is that analysts serve not only mutual funds or hedge funds but also other institutional investors. If they only served mutual funds or hedge funds, then one might not expect to see an effect on analyst coverage. To test this prediction, we utilize the same methodology as developed above for our analysis of institutional ownership. This methodology may also be viewed as an extension of earlier studies on the determinants of analyst coverage, such as Hong, Lim, and Stein (2000).

In this section, we estimate the following cross-sectional regression specification:

$$LOGCOV_{it} = b_0 + b_1\, SINDUM_{it} + \boldsymbol{b}_2\boldsymbol{X}_{it} + \varepsilon_{it}, \quad i = 1, \ldots, N, \tag{2}$$

where *SINDUM* equals one if the stock is a sin stock and zero otherwise; $\boldsymbol{X}_{it}$ is a vector of firm characteristics, and $\varepsilon_{it}$ is measurement error. The components of $\boldsymbol{X}_{it}$ and the estimation methodology are the same as described above for regression specification (1).

The results are presented in Panel C of Table 3, whose layout parallels that of Panel A above. Accordingly, we are able to quickly summarize our key findings. Notice that the coefficient in front of *SINDUM* is negative and statistically significant across four out of the six columns representing alternative specifications of our model. The coefficient in front of *GDUM* is tiny and statistically insignificant. In column 6, again our most conservative specification, the coefficient in front of *SINDUM* is −0.14 and is statistically significant at the 10% level of significance.[24] In terms of the absolute number of analysts, the typical firm has about 1.7 analysts covering it, whereas a sin stock has about 1.3 (about one-half fewer analysts or a 23% decline relative to the mean). As for the other variables, they attract coefficients that are similar to what already has been shown in Hong, Lim, and Stein (2000).

One potential alternative explanation for these findings is that sin stocks are mature companies that do little equity issuance and hence, attract little analyst coverage. To disentangle this effect, we construct from the SDC Equity Issuance Database the annual amount of equity issuances done by companies and include this as a control variable. Adding this variable does not change our results, and therefore we omit the results from this augmented specification for brevity.

In Panel D, we calculate the average coefficient for *SINDUM* using the specification from column 6 of Panel A for different sub-periods, 1980–1984, 1985–1989, 1990–1994, 1995–1999, and 2000–2006. In other words, rather than averaging across all the years which is how we got the coefficient in Panel A, we want to see how this coefficient has varied over these years. The idea is to see whether the shunning of sin stocks has been fairly stable over these years, as this will inform us as to the pricing implications we might expect from this neglect. Notice that the coefficient for 1980–1984 is −0.0660 while the coefficient at the end of the sample is of a comparable magnitude −0.0638. In the intervening periods, the coefficient bounces around these two values. So there does not appear to be a trend in terms of sin stocks being more shunned by institutions over this period.

These results are consistent with estimates from the Social Investment Forum, which produces reports on their estimates of the fraction of private portfolios that are screened based on these sin measures. Interestingly, going back to their 1990s reports, the fraction of screened dollars has been quite stable over the last 15 years. Importantly, they also establish similar figures for Europe.

In sum, we have obtained compelling evidence that sin stocks are less held or followed by certain institutions and analysts who discriminate against sin stocks for social norm rationales; and in their absence, arbitrageurs comprised of mutual funds and hedge funds (and individuals) are willing to buy these stocks. Moreover, these patterns have been fairly stable over the last two decades. These findings are consistent with both Becker's theory of racial discrimination by employers and Arrow's subsequent remark on arbitrage by employers who are willing to flout social convention to take advantage of unemployed, talented labor.

In our investigation of social norms in the stock market, the question that remains is whether social norms affect the prices of sin stocks, i.e., is there enough arbitrage capital to completely eliminate any norm-induced price effects?

### 4.3. Implications for stock prices

To address this question, we examine the return performance of sin stocks. If sin stocks are neglected and faced litigation risk heightened by social norms (in a stable manner) for the last few decades, as we claim, and there are limits to arbitrage, then we ought to find that sin stocks outperform comparables over this period.

We employ a number of methodologies to study the potential price effects of social norms on sin stocks. First, we employ the methodology of analyzing the time series of the returns of a sin stock portfolio, net of comparables, for evidence of any excess returns, after adjusting for various well-known predictors of stock returns such as the market portfolio. The first model we estimate is the CAPM:

$$EXCOMP_t = \alpha + \beta VWRF_t + \varepsilon_t, \quad t = 1, \ldots, T, \tag{3}$$

where $EXCOMP_t$ is the return on *SIN–COMP*, an equal-weighted portfolio long sin stocks and short their comparables; *VWRF* is the value-weighted market portfolio; and $\varepsilon_t$ represents a generic error term that is uncorrelated with all other independent variables. The coefficient of interest is $\alpha$, representing the excess return of the sin stock portfolio. We also consider three additional performance models by adding the portfolio returns *SMB*, *HML*, and *MOM* as independent variables in specification (3). Our most conservative specification, thus, is given by the following four-factor model:

$$\begin{aligned} EXCOMP_t = {} & \alpha + \beta_1 VWRF_t + \beta_2 SMB_t + \beta_3 HML_t \\ & + \beta_4 MOM_t + \varepsilon_t, \quad t = 1, \ldots, T, \end{aligned} \tag{4}$$

[24] In column 3, using *LOGMB* as a control rather than *PRINV* yields a larger point estimate on *SINDUM* (−0.1578), but weakens its statistical significance. The convention in the literature is to use *PRINV* which we follow here for comparability. When we include the other controls as in the later columns, then we still obtain a statistically significant estimate on *SINDUM* even when using *LOGMB* rather than *PRINV* as a control.

where $EXCOMP_t$ again denotes the return on the portfolio long sin stocks and short comparables, $\alpha$ is the excess return of that portfolio, $\beta_i$'s are loadings on the other portfolios that are used to predict $EXCOMP_t$, and $\varepsilon_t$ is a generic error term that is uncorrelated with all other independent variables. The loadings of our sin portfolio on the four factors are time varying. As such, we estimate these regressions on rolling three-year periods and then average the coefficients across these periods to obtain the results.

The results for the various specifications are presented in Panel A of Table 4. We first estimate the regression models using the period of 1965–2006. This is our benchmark period as we can obtain data to estimate cross-sectional return regression models and we know that all the stocks are steadily shunned during this sample. For completeness, we also estimate our results over the longer period, going back to 1926, though the caveat here is that tobacco has not always been considered a sin stock (see discussion in Section 4.7). The CAPM yields an alpha of 25 bps a month, which is statistically significant at the 10% level. For the two-factor model (market and SMB), the alpha is slightly higher at 30 bps a month with a 5% level of significance. When we subsequently include HML, the alpha goes down to 26 bps a month. When we add in MOM, our most conservative figure is 26 bps a month for the outperformance of sin stocks relative to their comparables, with significance at the 5% level. In Panel A, we also report these regression results estimated over the longer sample going back to 1926. Here, we find similar outperformance figures of around 30 bps a month and significance at the 5% level. So, in sum, our time-series results tell a consistent story that sin stocks outperform their comparables even after adjusting for standard factors in the literature.

We use cross-sectional variation to investigate whether sin stocks outperform other comparable stocks. We estimate the following return forecasting specification:

$$EXMRET_{it} = c_0 + c_1 SINDUM_{it-1} \boldsymbol{c}_2 \boldsymbol{X}_{it-1} + \varepsilon_{it}, \quad i = 1, \ldots, N, \tag{5}$$

where *EXMRET* is the return of stock *i*, net of the risk-free rate; *SINDUM* equals one if the stock is a sin stock and zero otherwise; $\boldsymbol{X}_{it-1}$ is a vector of firm characteristics; and $\varepsilon_{it}$ is measurement error. $\boldsymbol{X}_{it-1}$ includes various permutations of variables that have been defined earlier, including *LOGSIZE*, *RET*, and *LOGMB*. Since we now are estimating forecasting regression models, *BETA* represents the time-varying industry beta estimated using the past three years of monthly returns. Data on firm characteristics only begin in 1962. Accordingly, the sample for our estimation of Eq. (5) starts in 1965.

The coefficient of interest is $c_1$, which measures whether sin stocks have an abnormal return performance controlling for a host of other firm characteristics. The null hypothesis is that it is zero, whereas our prediction is that it will be significantly $>0$. $\boldsymbol{c}_2$ is the vector of loadings on the control variables. We then take the estimates from these monthly regressions and follow Fama and MacBeth (1973) in taking their time-series means and standard deviations, using Newey and West (1987) standard errors, to form our overall estimates of the effects of being a sin stock on return performance.

The results are presented in Panel B of Table 4. While we present various permutations of the regression specification (5), the main one to focus on is the specification described in column 6 that includes beta, size, past returns, past turnover, market-to-book, *GDUM*, and the log of the age of the firm (*LOGAGE*) as controls. *GDUM* picks up the effect due to comparables and *LOGAGE* has been shown to forecast firm profitability. In this specification, the coefficient in front of *SINDUM* is 0.0029 and is statistically significant at the 5% level of significance. In other words, sin stocks outperform other comparable stocks by about 30 basis points per month or about 3.5% per year. As earlier papers have also found, *BETA* is statistically insignificant in these cross-sectional regressions. *LOGSIZE* has a negative effect (big stocks underperform small stocks), past returns positively predict future returns (i.e., there is momentum over short horizons but not over long horizons), and turnover and market-to-book negatively forecast returns.

As we relax the controls, the coefficients get progressively weaker, though there is still statistical significance in column 4 (28 bps monthly excess performance at the 5% level of significance) when we control for size, market-to-book, momentum, and *GDUM* (the industry comparable dummy). Our results are telling us that one needs to adequately control for other characteristics that are correlated with a stock's sin status to capture a pure sin (neglect) effect. Controlling for *GDUM* (a comparables dummy) is especially important if we are to distinguish our sin effect from a more general industry effect. Here, the broader industries to which the sin stocks belong might have lower returns for other reasons. As such, we view the 28 bps figure obtained roughly from columns 4–6 with 5% level of significance as being the right measure of our sin effect on returns and being very statistically significant. There is also little change in the results from columns 4–6, which is telling us that these four characteristics (size, market-to-book, momentum, and *GDUM*) are all that is essential to estimate the pure sin effect. Moreover, these cross-sectional regression results are giving us a nearly identical answer as the time-series regressions in Panel A estimated over the same sample period.

Finally, we compare the valuation ratios of sin stocks to their counterparts using the following specification adapted from Hong, Kubik, and Stein (2008):

$$VALUTION_{it} = d_0 + d_1\, SINDIUM_{it} + d_2 \boldsymbol{X}_{it} + \varepsilon_{it}, \quad i = 1, \ldots, N, \tag{6}$$

where *VALUATION* can be either *LOGMB*, *LOGPE*, or *LOGPEBITDA*. *SINDUM* equals one if the stock is a sin stock and zero otherwise. $\boldsymbol{X}_{it}$ is a vector of firm characteristics including return-on-equity (*ROE*) as well as the next three year's *ROEs* (*FROE*, *F2ROE*, *F3ROE*), research and development expenditures as a fraction of sales (*RDSALES*), a dummy variable if the firm is missing R&D data (*RDMISS*), an S&P 500 dummy (*SP500*), and our *GDUM* dummy capturing sin stocks and sin stock comparables (food, soda, fun, meals, and hotels). $\varepsilon_{it}$ is measurement error.

**Table 4**
Price and return performance of sin stocks.

Panel A reports the average coefficients obtained from the time-series regressions of a portfolio (SIN–COMP) that is long *SIN* (the monthly return for an equal-weighted portfolio of sin stocks—alcohol, gaming, and tobacco) and short *COMP* (the monthly return for an equal-weighted portfolio of comparable stocks) on a host of well-known factors. Each regression is estimated using a 36-month window of data for the period of 1965–2006 as well as for the period of 1926–2006. *MKTPREM* is the excess monthly return of the value-weighted CRSP index. *SMB* is the return of a portfolio long small stocks and short large stocks. *HML* is the return of a portfolio long high book-to-market stocks and short low book-to-market stocks. *MOM* is the return of a portfolio long past 12-month return winners and short past 12-month return losers. Standard errors are adjusted for serial correlation using the Newey-West correction. Panel B reports the results of Fama and MacBeth (1973) cross-sectional regressions for the period of 1965–2006 of *EXMRET*, the monthly return of a stock net of the risk-free rate on the lagged (previous month) values of a set of well-known predictors of stock returns. *SINDUM* equals one if a stock is a sin stock and zero otherwise. *LOGSIZE*, *LOGMB*, and *RET* are now calculated at the end of each month. *BETA* is just the firm's industry beta, which is now calculated at the end of each month using the past 36-months of data. *TURN*, calculated at the end of month, is the daily share turnover during the past 12 months. *EXMRET* at month *t* is regressed on the previous month values of *LOGSIZE, LOGMB, RET, BETA, TURN*, and *BLEV* which are denoted by *LOGSIZE*1, *LOGMB*1, *RET*1, *BETA*1, *TURN*1, and *BLEV*1. *LOGAGE*1 is the log of the age of the company. *GDUM* equals one if a stock is a sin stock or comes from the Fama and French (1997) industry groupings: 2 (food), 3 (soda), 7 (fun), and 43 (meals), and zero otherwise. Panel C reports the results for a Fama-MacBeth OLS regression of the log of a stock's market-to-book ratio (*LOGMB*), log price-to-earnings ratio (*LOGPE*), and log price-to-EBITDA ratio (*LOGPEBITDA*) on a host of explanatory variables for the period 1965–2006. *SINDUM* equals one if a stock is a sin stock and zero otherwise. *ROE* is firm's return on equity, *FROE* is next year's ROE. *RDSALES* is the ratio of a firm's research and development expenditures to firm sales, *RDMISS* equals one if a firm is missing R&D expenditure data and zero otherwise. *SP500* equals one if the stock is in the S&P 500 index and zero otherwise. Column 1 of Panel D mirrors column 6 of Panel B, but the monthly returns are calculated without considering the three days around each earnings announcement. Column 2 additionally includes *BLEV*, and column 3 excludes tobacco and its comparable (food) from the specification. Panel E uses the same setting as Panel C, but the sin portfolio excludes the tobacco industry and its comparable (food). *** 1%; ** 5%; and * 10% significance.

Panel A: Time-series return regressions (net of comparables)

1965–2006

| | ALPHA | MKTPREM | SMB | HML | MOM |
|---|---|---|---|---|---|
| *SIN–COMP* | 0.0025* | −0.0060 | | | |
| | (0.0014) | (0.0399) | | | |
| *SIN–COMP* | 0.0030** | 0.0233 | −0.0944*** | | |
| | (0.0015) | (0.0341) | (0.0347) | | |
| *SIN–COMP* | 0.0026* | 0.0358 | −0.0993*** | 0.0832** | |
| | (0.0014) | (0.0365) | (0.0342) | (0.0410) | |
| *SIN–COMP* | 0.0026** | 0.0442 | −0.1058*** | 0.0733* | −0.0101 |
| | (0.0013) | (0.0401) | (0.0362) | (0.0447) | (0.0406) |

1926–2006

| | ALPHA | MKTPREM | SMB | HML | MOM |
|---|---|---|---|---|---|
| *SIN–COMP* | 0.0030** | −0.1397*** | | | |
| | (0.0015) | (0.0397) | | | |
| *SIN–COMP* | 0.0033** | −0.0912** | −0.2482*** | | |
| | (0.0014) | (0.0366) | (0.0621) | | |
| *SIN–COMP* | 0.0028** | −0.0325 | −0.2256*** | −0.0343 | |
| | (0.0014) | (0.0280) | (0.0587) | (0.0620) | |
| *SIN–COMP* | 0.0031** | −0.0190 | −0.2390*** | −0.0309 | 0.0063 |
| | (0.0014) | (0.0293) | (0.0607) | (0.0501) | (0.0410) |

Panel B: Cross-section regressions: 1965–2006

| Variable | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| *SINDUM* | 0.0021 | 0.0015 | 0.0018 | 0.0028** | 0.0028** | 0.0029** |
| | (0.0016) | (0.0015) | (0.0014) | (0.0013) | (0.0013) | (0.0013) |
| *LOGSIZE1* | −0.0026*** | −0.0015*** | −0.0014*** | −0.0014*** | −0.0015*** | −0.0013*** |
| | (0.0005) | (0.0005) | (0.0005) | (0.0005) | (0.0005) | (0.0005) |
| *LOGMB1* | | −0.0109*** | −0.0114*** | −0.0114*** | −0.0112*** | −0.0114*** |
| | | (0.0011) | (0.0009) | (0.0009) | (0.0009) | (0.0008) |
| *RET1* | | | 0.0411** | 0.0400*** | 0.0510*** | 0.0500*** |
| | | | (0.0192) | (0.0191) | (0.0179) | (0.0178) |
| *GDUM* | | | | −0.0019 | −0.0021* | −0.0020* |
| | | | | (0.0012) | (0.0011) | (0.0011) |
| *BETA1* | | | | | 0.0002 | 0.0001 |
| | | | | | (0.0020) | (0.0020) |
| *TURN1* | | | | | −0.0368 | −0.0465 |
| | | | | | (0.1930) | (0.1889) |
| *LOGAGE1* | | | | | | −0.0006 |
| | | | | | | (0.0005) |

Table 4 (*continued*)

Panel C: Valuation regressions: 1965–2006

| Variable | *LOGMB* | *LOGPE* | *LOGPEBITDA* |
|---|---|---|---|
| *SINDUM* | −0.1540*** | −0.1508*** | −0.2035*** |
| | (0.0136) | (0.0220) | (0.0452) |
| *ROE* | 0.7834** | −0.8027*** | −0.4454*** |
| | (0.3132) | (0.1097) | (0.1089) |
| *RDSALES* | 4.7701*** | 5.9278*** | 7.9579*** |
| | (1.4821) | (1.2760) | (0.8626) |
| *RDMISS* | −0.0793 | −0.0048 | 0.0278 |
| | (0.0604) | (0.0386) | (0.0328) |
| *SP500* | 0.2486*** | 0.0972** | 0.0516 |
| | (0.0626) | (0.0464) | (0.0435) |
| *GDUM* | 0.0944*** | 0.1031*** | 0.0746*** |
| | (0.0274) | (0.0241) | (0.0253) |
| *FROE* | 0.4542** | 0.0216 | −0.0336 |
| | (0.1986) | (0.0790) | (0.0311) |
| *F2ROE* | 0.1318*** | −0.0454 | −0.0149 |
| | (0.0482) | (0.0524) | (0.0440) |
| *F3ROE* | 0.0937 | −0.0619 | −0.1312* |
| | (0.0655) | (0.0509) | (0.0696) |

Panel D: Additional cross-section regressions: 1965–2006

| *Variable* | Excl. announcement return | Book leverage | Excluding tobacco |
|---|---|---|---|
| *SINDUM* | 0.0030** | 0.0032** | 0.0021* |
| | (0.0012) | (0.0013) | (0.0013) |
| *LOGSIZE1* | −0.0012** | −0.0013*** | −0.0013*** |
| | (0.0005) | (0.0005) | (0.0005) |
| *LOGMB1* | −0.0094*** | −0.0115*** | −0.0114*** |
| | (0.0007) | (0.0008) | (0.0008) |
| *RET1* | 0.0223 | 0.0494*** | 0.0500*** |
| | (0.0170) | (0.0181) | (0.0177) |
| *GDUM* | −0.0020** | −0.0019* | −0.0018 |
| | (0.0010) | (0.0011) | (0.0013) |
| *BETA1* | −0.0004 | −0.0001 | 0.0002 |
| | (0.0018) | (0.0019) | (0.0020) |
| *TURN1* | 0.2057 | 0.0358 | −0.0488 |
| | (0.1797) | (0.2019) | (0.1893) |
| *LOGAGE1* | −0.0005 | −0.0007 | −0.0006 |
| | (0.0004) | (0.0005) | (0.0005) |
| *BLEV1* | | −0.0027* | |
| | | (0.0014) | |

Panel E: Valuation regressions excluding tobacco: 1965–2006

| Variable | *LOGMB* | *LOGPE* | *LOGPEBITDA* |
|---|---|---|---|
| *SINDUM* | −0.1652*** | −0.1212*** | −0.2044*** |
| | (0.0325) | (0.0414) | (0.0495) |
| *ROE* | 0.7838** | −0.8068*** | −0.4463*** |
| | (0.3132) | (0.1102) | (0.1096) |
| *RDSALES* | 4.7811*** | 5.9382*** | 8.0037*** |
| | (1.4858) | (1.2802) | (0.8759) |
| *RDMISS* | −0.0809 | −0.0068 | 0.0266 |
| | (0.0595) | (0.0375) | (0.0319) |
| *SP500* | 0.2470*** | 0.0955** | 0.0489 |
| | (0.0626) | (0.0467) | (0.0432) |
| *GDUM* | 0.1531*** | 0.1776*** | 0.1578*** |
| | (0.0441) | (0.0390) | (0.0280) |
| *FROE* | 0.4558** | 0.0241 | −0.0324 |
| | (0.1991) | (0.0781) | (0.0317) |
| *F2ROE* | 0.1334*** | −0.0424 | −0.0125 |
| | (0.0486) | (0.0512) | (0.0446) |
| *F3ROE* | 0.0941 | −0.0610 | −0.1306* |
| | (0.0658) | (0.0509) | (0.0700) |

The results are presented in Panel C of Table 4. In column 1, we estimate the cross-sectional regression year by year for *LOGMB*. The coefficients are subsequently aggregated using the methodology of Fama and MacBeth (1973), in which standard errors are adjusted for possible autocorrelation. The coefficient on *SINDUM* is −0.1540 and is significant at the 1% level of significance. We thus find sin stocks' market-to-book ratios are smaller relative to comparables by nearly 15.4%. The coefficients on *ROE*, *RDSALES*, and *SP500* are positive and significant, while the coefficient on the dummy variable for missing R&D data is negative and insignificant. These latter findings are consistent with the results reported by Hong, Kubik, and Stein (2008). Interestingly, the coefficient on *GDUM* is positive, suggesting that the prices of sin stock comparables actually are higher than those of other stocks.

In column 2, we estimate the same model, this time using *LOGPE*. The coefficient in front of *SINDUM* drops slightly to −0.1508, but it is still significant at the 1% level of significance. The other coefficients remain qualitatively similar. Importantly, the coefficient in front of *SINDUM* for *LOGEBITDA* is −0.2035 and significant at the 1% level. In summary, all three methodologies that we have used to investigate the price effects of social norms on sin stocks yield economically interesting effects.

### 4.4. Robustness checks

In Panel D, we consider some robustness checks. We take the regression specification in column 6 of Panel B and reestimate it by netting out returns around earnings announcement dates (defined as the three-day return around quarterly earnings announcements). The result is in column 1 of Panel D. We find a negligible change. The worry here is that the results might be due to unexpected earnings news, which it does not appear to be. We next worry that our result might be due to firm leverage since sin companies tend to be more levered as we show below. We insert as an additional control variable book leverage and do not find a significant difference. Also, controlling for market leverage yields similar results to those in the second column. In the third column, we omit tobacco from our sin stocks. The worry here is that tobacco might have done unexpectedly well in evading all their litigation risk during this period and hence might be driving all our results. We find that the sin portfolio, net of tobacco this period, yields an alpha of 21 bps and significance at the 10% level. The results are slightly lower than those in column 6 of Panel B but are not significantly different from the 29 bps figure, though it does suggest that tobacco's litigation risk this period might have contributed to a slightly higher than sin portfolio return. One caveat is that even this might not be conservative enough since one could also argue that the gaming industry might have unexpectedly good news. But it is difficult to eliminate both gaming and tobacco industries and still get statistical significance. Our additional analyses below provide comfort as to the robustness of our main findings.

In Panel E, we consider a robustness check of Panel C in that we reestimate it without tobacco stocks. We find similar results. For *LOGMB*, the coefficient in front of *SINDUM* is slightly higher at −0.1652 with significance at the 1% level. For *LOGPE*, the coefficient in front of *SINDUM* is slightly lower at −0.1212 but it is nonetheless very significant. For *LOGEBITDA*, the coefficient is slightly higher at −0.2044 and again very significant. In sum, all of our return and valuation regressions are robust.

We have explored a number of other robustness checks to the findings reported in this paper, but omit their full reporting for brevity. In particular, as we mentioned earlier in Section 2, defense stocks are considered by some to be sinful. We have redone all of our analyses with defense stocks included in the sin category. Our results remain qualitatively similar. We nevertheless focus on the so-called Triumvirate of Sin (alcohol, tobacco, and gaming) since there is much stronger consensus on the characterization of these industries.

### 4.5. Reconciling magnitudes from return and valuation regressions

In this section, we analyze whether our return results (Panels A, B, and D of Table 4) are consistent with our valuation or "Q" effects (Panels C and E). The question is whether the return results are consistent with the implied-return figure from valuation regressions (after taking into account estimation error from these two sets of regressions). First, note the range of return estimates. The baseline figure from the time-series regression (Panel A) is 26 bps or 3.1% a year. The one from the cross-sectional regression (Panel B) is 28 bps or 3.4% a year. Now, this figure might be influenced by unexpected good litigation news for tobacco stocks. The result from Panel D suggests that a more conservative number is 21 bps or 2.5% a year. But the range is from 2.5% to 3.4% a year. The 21 bps per month figure is perhaps more reasonable and in line with the size effect that is the basis behind the neglect model of Merton (1987). Indeed, as we show in Section 4.8 on international stocks, we get a 21 bps or 2.5% a year extra return from international stocks. So the figure of 2.5% is a good one to focus on for the purposes of checking the reasonableness of the valuation regression magnitudes.

There are different reasons to like each of the three valuation measures but they all deliver similar magnitudes. From Panel C, the low end in terms of a price effect is around 15% using *LOGPE* and the high end is a 20% price effect using *LOGPEBITDA*. Indeed, the estimation error of the price-to-EBITDA measure is such that it implies as much as a 25% lower valuation. To relate these valuation ratio results to our return results we use the Gordon growth formula, $1/(r-g)$, where $r$ is the discount rate and $g$ is the growth rate. As a benchmark, let us assume $r$ is 12% for a typical stock and $g$ is 4% (which is the mid-point of the range of values that we typically observe in the valuation literature). We can ask how much of an increase in $r$ is required to get a 17% lower valuation for sin stocks. At these parameters, this would imply around a 2% excess return, which is not far away from the 2.4% figure that we calculated. Indeed, if we take into account estimation error, an even higher figure of 3.4% from the return regressions would fall in the range of return estimates implied from the price-to-EBITDA regressions.

### 4.6. Calibrations with Merton model

The next question one can raise then is whether these magnitudes can be roughly reconciled within a Merton-style segmented market model. First, a realistic calibration of a Merton-type model is done in Petajisto (2009). The pricing in his model is done by arbitrageurs or mutual funds. Managers have to pay a fee to cover a stock and uncover mispricings, which allows them to earn their excess returns to justify their fees. Petajisto uses the realistic fees charged by mutual funds (analog to Merton's fixed cost) of 1–1.5% of assets under management, as a means to figure out how big of a price effect we should see assuming a pure 10% supply shock in the stock (either positive or negative). He is interested in seeing whether one can obtain realistic S&P 500 inclusion effects in which when a company gets included into the S&P 500, one sees a price increase for some stocks of as much as 15%. This example is very similar to our sin stocks situation except that there is a positive supply shock (with certain large passive institutions staying away from these stocks for political or ethical reasons) as opposed to a negative supply shock (passive index funds buying shares of the included company). The supply shocks in either situation are similar in that estimates place the fraction of dollars under management that undergo an ethical screen at 1-in-9 to 1-in-10 dollars, very similar to estimates for S&P 500 passive index funds. Petajisto shows that under realistic parameters for fees, volatility, and persistence, we can get 15% pricing effects—which is in the ball park of the numbers we find for sin stocks.

Moreover, we present a simplified version of Heinkel, Kraus, and Zechner (2001), henceforth HKZ, who develop a model to consider the price implications of ethical investing that excludes companies that pollute. We show in Appendix A that this model can generate realistic pricing effects as well. First, in HKZ, they allow for the possibility of a reformed company—an unacceptable company that pays a fixed cost to become an acceptable firm. This does not apply to our discussion of sin stocks since they cannot really change who they are by paying a fixed cost. As such, we simplify the HKZ model by allowing for only two types of firms, acceptable and unacceptable (which would be the sin stocks). Second, as in HKZ, we assume that green investors only invest in acceptable companies, whereas neutral investors will invest in any company. Third, in HKZ, they assume that green investors and neutral investors have the same risk aversion. In reality, institutions which follow screening rules tend to be much less risk averse than individual/retail investors. In our set-up, we allow for the green investors to be more risk tolerant than the neutral investors. We then show that we can get realistic pricing differences consistent with the data even assuming reasonable parameters on the fraction of green investors (say 10% or 15%) and the fraction of unacceptable companies (say 5%). The reason is that if green investors are massive institutions such as CalPERS (which we describe below) with long horizons and deep pockets, and neutral investors for sin companies are likely to be individual investors who are far more risk averse, then one can still get some realistic differences since the holders of the sin stocks will tend to be much more risk-averse individuals.

### 4.7. Further tests

We now test some auxiliary implications of our hypothesis. First, we investigate the implications of the price effects of social norms for the financing decisions of sin companies. If the equity of sin stocks is undervalued because of limited risk sharing or investors misestimating the risk of these companies as we have argued in this paper, and hence it is more expensive for sin companies to finance operations using equity, then we should expect them to use private debt to finance their operations. Debt markets offer the additional advantage of being less transparent than equity markets. Whereas mutual funds and institutions are required to disclose their positions in equities on a semi-annual basis, no such requirements exist for investments in corporate bonds. Accordingly, it is difficult to figure out for a given company who its financiers are on the debt side. While it is possible to track down large public issuances of corporate bonds, large amounts of bank debt are difficult to trace. Former SEC chairman Arthur Levitt, who pushed to increase transparency in the bond market during his tenure, once remarked, "The sad truth is that investors in the corporate bond market do not enjoy the same access to information as a car buyer or a home buyer or, dare I say, a fruit buyer. And that's unacceptable".[25]

To determine if sin companies indeed use the debt market more than the equity market, we implement a standard cross-sectional regression specification used in the corporate finance literature to explain capital structure, i.e., leverage (e.g., Baker and Wurgler, 2002). We estimate the following cross-sectional regression specification:

$$\mathit{FinancingDecision}_{it} = e_0 + e_1\, \mathit{SINDUM}_{it} + \boldsymbol{e}_2 + \boldsymbol{X}_{it} + \varepsilon_{it}, \quad i = 1, \ldots, N \tag{7}$$

where *SINDUM* equals one if the stock is a sin stock and zero otherwise; $\boldsymbol{X}_{it}$ is a vector of firm characteristics; and $\varepsilon_{it}$ is measurement error. $\boldsymbol{X}_{it}$ includes the variables *GDUM*, *TOBQ*, *TANG*, *PROFIT*, and *LOGSALES*. The coefficient of interest is $e_1$, which measures whether sin firms have a different capital structure as compared to other companies. The null hypothesis is that it is zero, whereas our prediction is that it will be significantly $>0$ when a measure of leverage is used as the dependent variable in the regression. $\boldsymbol{e}_2$ is the vector of loadings on the control variables. We then take the estimates from a pooled

[25] The full text of Arthur Levitt's speech from which this quote has been excerpted can be downloaded from ⟨http://www.sec.gov/news/speech/speecharchive/1998/spch218.htm⟩.

**Table 5**
Corporate financing decisions of sin companies.

This table reports estimates from the pooled regression of sin companies' corporate financing decisions on well-known predictors of capital structure for the period of 1965–2006. The dependent variables are market leverage (*MLEV*), book leverage (*BLEV*), cash balances (*CASH*), overall firm payout ratio (*PAYOUT*), dividend payout ratio (*DIVPAY*), and share repurchases (*REP*). *SINDUM* equals one if a stock is a sin stock (alcohol, gaming, and tobacco) and zero otherwise. *GDUM* equals one if a stock is a sin stock or comes from the Fama and French (1997) industry groupings: 2 (food), 3 (soda), 7 (fun), and 43 (meals), and zero otherwise. *TOBQ* is Tobin's *Q*, which is market value of equity plus assets minus book value of equity over assets. *TANG* is firm tangibility defined as net plant, property, and equipment divided by total assets. *PROFIT* is earnings before interest, taxes, and depreciation divided by total assets. *LOGSALES* is the log of net sales. All variables have been winsorized at the 1% level. All regressions include one-digit SIC industry fixed effects. *** 1% significance; ** 5% significance; and * 10% significance.

| | *MLEV* | *BLEV* | *CASH* | *PAYOUT* | *DIVPAY* | *REP* |
|---|---|---|---|---|---|---|
| *SINDUM* | 0.0520** | 0.0688 | −0.0187 | −0.0056 | 0.0141 | −0.0148 |
| | (0.0244) | (0.0445) | (0.0243) | (0.0911) | (0.0829) | (0.0235) |
| *GDUM* | −0.0073 | 0.0031 | 0.0216** | 0.0147 | −0.0351 | 0.0452** |
| | (0.0130) | (0.0098) | (0.0102) | (0.0620) | (0.0568) | (0.0218) |
| *TOBQ* | −0.0370*** | −0.0191*** | 0.0435*** | −0.0071*** | −0.0057*** | 0.0001 |
| | (0.0053) | (0.0024) | (0.0018) | (0.0025) | (0.0020) | (0.0008) |
| *TANG* | 0.2694*** | 0.3212*** | −0.2852*** | 0.1808* | 0.2461** | −0.0550* |
| | (0.0438) | (0.0343) | (0.0324) | (0.0923) | (0.1209) | (0.0326) |
| *PROFIT* | −0.1342*** | −0.1830*** | −0.0420 | 0.2777*** | 0.1362*** | 0.1330*** |
| | (0.0369) | (0.0318) | (0.0335) | (0.0375) | (0.0231) | (0.0174) |
| *LOGSALES* | 0.0133 | 0.0162* | −0.0564*** | 0.0145 | 0.0040 | 0.0074* |
| | (0.0085) | (0.0093) | (0.0103) | (0.0129) | (0.0125) | (0.0038) |

regression and cluster standard errors by the 49 Fama and French (plus gaming) industries.

The results are presented in Table 5. In column 1, the dependent variable is market leverage, *MLEV*. The coefficient in front of *SINDUM* is positive (0.0520) and statistically significant at the 5% level of significance. The typical company has a market leverage of 0.27, and thus, we find that a sin company has a 19.3% higher leverage ratio than the typical company. The coefficient on *GDUM* is negative but it is not statistically significant. The coefficients in front of the other variables (other than *GDUM* and *LOGSALES*) are significant, as found in earlier work. In the regression that uses book leverage (*BLEV*) as the dependent variable, the coefficient on *SINDUM* again comes in with the right sign, but it is not as statistically significant as its counterpart in the regression analysis of *MLEV*. For completeness, we also look to see whether there are differences in other financing decision variables (*CASH*, *PAYOUT*, *DIVPAY*, *REP*) between sin stocks and other companies. The sin stock effects are not statistically significant in these other regressions. In sum, it appears that sin companies rely more on the debt market for financing than do other firms, consistent with the hypothesis that they face a disadvantage in the equity market, but they do not differ markedly in their other financing decisions.

We also attempted to test indirect implications of our hypotheses. First, if sin stocks are more closely held, then the number of shares available to investors could decrease and cause similar patterns to the ones we describe. Of course, sin effect could also be a possible explanation of these increased close holdings, but likely not unique. We explore this possibility by relating the degree of close ownership, measured by the non-float fraction of outstanding shares, to the sin characteristic of each security. Our evidence indicates a weak negative relationship between sin stocks and close ownership. The results are similar if we use an alternative measure of ownership based on large-level institutional ownership. Hence, it is unlikely that the close-ownership effect drives our results. Second, since sin stocks are followed less by analysts and institutions, they may have higher illiquidity or bid–ask spreads. However, we find only negligible differences in bid–ask spreads between sin stocks and other stocks. There are a few reasons that may explain this negative finding. First, most sin companies are highly regulated and have much better disclosure than do other companies. Accordingly, there may be less information asymmetry for these stocks, leading to potentially lower spreads. Second, the lack of institutional ownership actually has an ambiguous effect on spreads depending on assumptions about the trading process. Therefore, indirect implications regarding liquidity or other market microstructure aspects of these companies are difficult to measure unless we have more direct measures of the informational environments of companies.

In addition, we also attempted to exploit the time variation in the social norms governing tobacco as a means to test our hypothesis. As we described earlier in Section 2, tobacco was not widely considered to be sinful until the mid-1960s. Accordingly, our hypothesis suggests that before that time, tobacco stocks should have been held by institutions and covered by analysts as frequently as other stocks, and they should not have earned excess returns. Unfortunately, due to data limitations, we can only test the implications regarding returns, since data on ownership and coverage are only available since the 1980s.

One prediction is that we should see tobacco stocks underperform over the period of the late 1940s (when anecdotal evidence suggested the change in norms with previous bad reports about health in the late 1940s) until the mid-1960s (when essentially even the government recognized that tobacco was bad and imposed many restrictions). The problem is finding the exact right dates to test this period of transition. The choices are arbitrary but not that sensitive to within a few years. In particular, our initial date is 1947 when the first really bad reports started coming out about tobacco and our ending date is 1965 when Congress passed important legislation regulating the advertising of cigarettes. One thing to note is that tobacco companies fought back in the late 1950s with false claims about filtered cigarettes that they prevented lung cancer. Although this action temporarily boosted the status of tobacco companies, these claims were rejected by the early 1960s and the transition of tobacco to being a sin stock was complete. We perform this exercise and the results are available from the authors. In sum, we find that tobacco underperformed the market by a significant 3% a year or something on the order of 40% over the period 1947–1965. This is a huge drop in prices in light of the generally good performance of tobacco throughout its history. This evidence has to be caveated with the disclaimer that testing these long trends is problematic for many reasons including changes in fundamentals throughout the periods. However, the evidence is largely consistent with there being a change in norms during this period of the 1950s–1960s.

### 4.8. *International stocks*

We extend our analysis to international markets: Canada, France, Germany, Italy, Netherlands, Spain, Switzerland, and United Kingdom. The data for these countries from Datastream only cover the period of 1985–2006 and we are only able to obtain a subset of the control variables that we had in the US regression. However, we are able to get the key characteristics of firm size, book-to-market, past returns, and firm comparables. As a result, we can run an analog to the key baseline regression we had in Panel B column 4 of Table 4, enhanced with the country fixed-effects.

Various surveys from the Social Investment Forum indicate that these markets have very similar attitudes toward sin stocks as does the US. One key difference though is that tobacco did not face nearly the same intense litigation as in the US. Also, institutions in these countries have had an aversion to sin stocks for quite some time. The key point of this analysis is to show that the US results are not spurious. To the extent that we find significant effects for the international sample, this is an out-of-sample test for our paper and will provide a high level of confidence in the robustness of our US results. One

**Table 6**
International evidence.

Panel A reports summary statistics about the sin stocks for the sample of international companies. We report country-by-country the number of sin stocks that fall into the three subgroups of tobacco, alcohol, and gaming. Panel B reports the results of the Fama and MacBeth (1973) cross-sectional regressions for the period of 1985–2006 of *EXMRET*, the monthly return of a stock net of the risk-free rate on the lagged (previous month) values of a set of well-known predictors of stock returns. *SINDUM* equals one if a stock is a sin stock and zero otherwise. *LOGSIZE*, *LOGMB*, and *RET* are calculated at the end of each month. *EXMRET* at month $t$ is regressed on the previous month values of *LOGSIZE*, *LOGMB*, and *RET*. *GDUM* equals one if a stock is a sin stock or comes from the Fama and French (1997) industry groupings: 2 (food), 3 (soda), 7 (fun), and 43 (meals), and zero otherwise. All regressions include country fixed effects and standard errors are adjusted for serial correlation as in Newey and West (1987). *** 1% significance; ** 5% significance; and * 10% significance.

Panel A: Distribution of sin stocks by country: 1985–2006

| Country | Number of sin stocks |
|---|---|
| Canada | 14 |
| France | 14 |
| Germany | 31 |
| Italy | 2 |
| Netherlands | 4 |
| Spain | 10 |
| Switzerland | 3 |
| UK | 34 |
| Total | 112 |

Panel B: OLS cross-sectional regressions for returns: 1985–2006

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| *SINDUM* | 0.0015 | 0.0014** | 0.0013** | 0.0021* |
| | (0.0010) | (0.0006) | (0.0004) | (0.0013) |
| *LOGSIZE1* | −0.0010 | −0.0004 | −0.0006 | −0.0006 |
| | (0.0008) | (0.0007) | (0.0006) | (0.0006) |
| *LOGMB1* | | −0.0023*** | −0.0040*** | −0.0040*** |
| | | (0.0005) | (0.0006) | (0.0006) |
| *RET1* | | | 0.0115*** | 0.0115*** |
| | | | (0.0010) | (0.0010) |
| *GDUM* | | | | −0.0008 |
| | | | | (0.0012) |
| Country fixed effects | Yes | Yes | Yes | Yes |

thing to keep in mind is that the international sample is much shorter than is that of the US and so the statistical significance will naturally be lower.

Panel A of Table 6 gives a breakdown of the number of sin stocks by country where sin is defined similarly as tobacco, alcohol, and gaming. Panel B shows the results of the cross-sectional regression. Notice that the results are both statistically and economically significant. Taking the most elaborate specification, column 4 of Table 6, we find that sin stocks outperform by 21 bps a month or about 2.5% per year. This effect is economically and statistically significant at the 10% level. The statistical significance is comforting given the small sample and the economic magnitude is also reassuring. The sin return estimate from the international sample (21 bps a month) is smaller than that obtained from the comparable US estimation (28 bps a month) but this difference is not statistically significant.

However, it is possible that this difference might be due to the fact that tobacco firms did not face nearly the same litigation pressure abroad as they did in the US. This litigation risk and the subsequent good outcomes for US firms in light of this litigation likely contributed to the higher returns of tobacco stocks in the US. This is consistent with the robustness checks in the US sample in which we estimate the sin stock returns excluding tobacco and also find a 21 bps monthly effect for non-tobacco sin stocks. But again, all these estimates are very close to each other in terms of economic and statistical significance and when taken in totality they provide strong proof of the robustness of our sin return effect.

## 5. Conclusion

In this paper, we provide evidence of significant effects of social norms on markets by studying the investing environment of "sin" stocks—publicly traded companies involved in the production of alcohol, tobacco, and gambling. In particular, we show that there is a significant price effect on the order of 15–20% from large institutional investors shunning sin stocks. Our paper has important implications on the emerging literature on socially responsible investing. A key issue in this literature is whether socially responsible investing makes a difference in terms of getting corporations to change their behavior. Or is socially responsible investing simply about feeling good, with little consequences for real investments? Our paper takes an important first step in answering this question in that we show that at least for sin stocks, the neglect of these stocks by large institutions affected their cost of capital in a significant way. Unfortunately, there is little in the way of change that these companies can make in contrast to say, companies that pollute. Hence, an important agenda for future research is to see to what extent socially responsible investing can affect the cost of capital for a broader range of companies and to connect changes in the behavior of these companies to distinct economic channels such as the cost of capital.

## Appendix A

This appendix provides details on the calibration of a simplified model based on HKZ. In brief, we consider a one-period model with $N$ firms of the following two types: $N_A$ acceptable ($A$) firms that satisfy the investing criteria of the green investors and $N_P$ unacceptable (polluting) ($P$) firms that do not satisfy such criteria. Each $A$ firm uses the clean technology and generates a normally distributed cash flow with mean $\mu_A$ and variance $\sigma_A^2$. Firm $P$ uses the polluting technology and generates normally distributed cash flow with mean $\mu_P$ and variance $\sigma_P^2$. The cash flows of firms $A$ are perfectly correlated as are the cash flows of firms $P$; the covariance between $A$ and $P$ technology is $\sigma_{AP}$. In addition, a riskless asset with a rate of return normalized to zero is available in perfectly elastic supply. Borrowing is allowed, but short selling is prohibited.

In this economy, there are two types of investors, $g$ (*green*) and $n$ (*neutral*), who differ in terms of their attitude to environmental damage. The former group abstains from holding shares in unacceptable firms, while the latter group does not internalize this fact in their preferences. In the context of our paper, the $g$ investors are the ones who are subject to social norms and do not invest in sin stocks, while the $n$ investors do not consider sin to be unacceptable. The total number of investors in the economy is $I$ with $I_g$ of green and $I_n$ of neutral investors. Green investors exhibit CARA preferences with risk tolerance $\tau_g$ and neutral investors have a similar preference but with a risk tolerance of $\tau_n$. Let $\varphi = \sigma_A^2\sigma_P^2 - \sigma_{AP}^2$. It is easy to show that the equilibrium prices for the two types of companies, $A$ and $P$, are given by the following two equations. The price of the $A$ firm is

$$P_A = \mu_A - 1(I_n\tau_n + I_g\tau_g)[N_P\sigma_{AP} + N_A\sigma_A^2], \tag{A.1}$$

and the price of the $P$ firm is given by

$$P_P = \mu_P - 1(I_n\tau_n + I_g\tau_g)[N_A\sigma_{AP} + N_P\sigma_P^2 + N_P(I_g/I_n)(\varphi/\sigma_A^2)(\tau_g/\tau_n)]. \tag{A.2}$$

Our goal is to verify whether prices observed for sin stocks could be obtained within the modeling framework outlined above. Absent these social norms, the assumptions of HKZ would yield the CAPM. However, CAPM is not valid with substantial ethical investors in the economy and idiosyncratic (or cash-flow) risk is priced. Empirical evidence suggests that the proportion of investors shunning sin stocks in the investing world approximately equals 10–15% and sin stocks represent about 5% of the stocks in the market. Therefore, with this number in mind we should expect the abnormal returns we obtain for sin stocks to be comparable to those calibrated to the model.

We begin our calibration with the parameters that are aimed to produce reasonable expected rates of returns. With these numbers we are able to generate the equity premium of about 8%, which is close to the excess market return shown in the literature. In addition, the variance–covariance matrix of cash flows produces reasonable results for the standard deviation of rate of return. In particular, we use the following set of parameters:

Mean cash flows: $\mu_P = \mu_A = 35$
Standard deviation of cash flows: $\sigma_P = 20$, $\sigma_A = 10$
Covariance of cash flows: $\sigma_{AP} = 100$
Total number of investors: $I = 1$

Total number of firms: $N = 1$
Number of firms with each technology: $N_P = 0.05$, $N_A = 0.95$
Aggregate risk tolerance: $\tau_n = 30$, $\tau_g = 100$

The quantity of our interest is abnormal return, which in the context of their model is equal to $(\mu_P/P_P){-}1$ (return of $P$ firms) compared to $(\mu_A/P_A){-}1$ (return of $A$ firms). HKZ assume that the cash-flow risk of all firms in the economy is the same (and equal to 10). However, in our context, sin stocks may have higher cash-flow risk than comparable firms because of the litigation risk that affects their operations. Since firm-specific risk is a priced risk in this framework, our calibration attempts to illustrate the sensitivity of the sin stock premium also conditioning on changing cash-flow risk. We use a figure of 20% to illustrate our calculation. At these numbers, the unacceptable firms have a higher return than acceptable firms by about 2% a year. The key parameters are that the green investors (presumably large institutions like CalPERS) are more risk tolerant than neutral investors who are more likely to be individual investors.

## References


Ahrens, D., 2004. Investing in Vice. St. Martin's Press, New York.
Akerlof, G., 1980. A theory of social custom, of which unemployment may be one consequence. Quarterly Journal of Economics 94, 749–775.
Altonji, J., Blank, R., 1999. Race and gender in the labor market. In: Card, D., Ashenfelter, O. (Eds.), Handbook of Labor Economics, vol. 3c. North-Holland, Amsterdam, pp. 3143–3259.
Arrow, K., 1972. Models of job discrimination. In: Pascal, A. (Ed.), Racial Discrimination in Economic Life. D.C. Heath, Lexington, MA, pp. 83–102.
Baker, M., Wurgler, J., 2002. Market timing and capital structure. Journal of Finance 57, 1–32.
Becker, G., 1957. The Economics of Discrimination. University of Chicago Press, Chicago.
Berman, D., 2002. Why sin is good: tobacco, alcohol, and gaming stocks can add sizzle to your portfolio? Money Sense, November.
Case, A., Katz, L., 1991. The company you keep: the effects of family and neighborhood on disadvantaged youths. NBER Working Paper no. 3705.
Chen, D., Bin, F., 2001. Effects of legislation events on US gaming stock returns and market turnings. Tourism Management 22, 539–549.
Del Guercio, D., 1996. The distorting effect of the prudent man law on institutional equity investments. Journal of Financial Economics 40, 31–62.
Elster, J., 1989. Social norms and economic theory. Journal of Economic Perspectives 3, 99–117.
Falkenstein, E., 1996. Preferences for stock characteristics as revealed by mutual fund portfolio holdings. Journal of Finance 51, 111–135.
Fama, E., French, K., 1997. Industry costs of equity. Journal of Financial Economics 43, 153–193.
Fama, E., MacBeth, J., 1973. Risk, return, and equilibrium: empirical test. Journal of Political Economy 81, 607–636.
Geczy, C., Stambaugh, R., Levin, D., 2003. Investing in socially responsible mutual funds. Unpublished Working Paper, University of Pennsylvania, Wharton.
Glaeser, E., Sacerdote, B., Scheinkman, J., 1996. Crime and social interactions. Quarterly Journal of Economics 111, 502–548.
Glaeser, E., Scheinkman, J., 2003. Non-market interactions. In: Dewatripont, M., Hansen, L., Turnovsky, S. (Eds.), Advances in Economics and Econometrics: Theory and Applications, Eight World Congress. Cambridge University Press, Cambridge, pp. 339–370.
Gompers, P., Metrick, A., 2001. Institutional investors and equity prices. Quarterly Journal of Economics 116, 229–259.
Greene, W., 2000. Econometric Analysis. Macmillan, New York.
Heinkel, R., Kraus, A., Zechner, J., 2001. The effect of green investment on corporate behavior. Journal of Financial and Quantitative Analysis 35, 431–449.
Hong, H., Kubik, J., Stein, J., 2004. Social interaction and stock market participation. Journal of Finance 59, 137–163.
Hong, H., Kubik, J., Stein, J., 2008. The only game in town: stock price consequences of local bias. Journal of Financial Economics 90, 20–37.
Hong, H., Lim, T., Stein, J., 2000. Bad news travels slowly: size, analyst coverage and the profitability of momentum strategies. Journal of Finance 55, 265–295.
Levitt, S., 2004. Testing theories of discrimination: evidence from weakest link. Journal of Law and Economics 47, 431–452.
Merton, R., 1987. A simple model of capital market equilibrium with incomplete information. Journal of Finance 42, 483–510.
Moulton, B., 1986. Random group effects and the precision of regression estimates. Journal of Econometrics 32, 385–397.
Newey, W., West, K., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Petajisto, A., 2009. Why do demand curves for stocks slope down? Journal of Financial and Quantitative Analysis, forthcoming.
Romer, D., 1984. The theory of social custom: a modification and some extensions. Quarterly Journal of Economics 99, 717–727.
Shleifer, A., Vishny, R., 1997. The limits of arbitrage. Journal of Finance 52, 35–55.
Teoh, S., Welch, I., Wazzan, P., 1999. The effect of socially activist investment policies on the financial markets: evidence from the South African Boycott. Journal of Business 72, 35–89.
Waxler, C., 2004. Stocking Up on Sin. John Wiley and Sons, Inc., New Jersey.