# Prosocial Norms and the Value of Aligned Values: Evidence from Over a Century

Joel F. Houston Sehoon Kim Boyuan Li *

This draft: November, 2025

**Abstract**

History is punctuated by changes in social norms about the responsibilities expected of businesses. Applying state-of-the-art natural language processing techniques to U.S. historical news text, we develop a high-frequency time-series index dating back to the late 19th century that measures public attention and social norms about environmental and social (E&S) issues around businesses. We explore the evolution of context captured by the Prosocial Norm Index (PNI) and its relation to macroeconomic fluctuations, asset prices, and corporate decisions. Going from periods of macroeconomic and social instability to times of relative prosperity, we observe a migration of public attention and concern from social to environmental issues. At the firm-level, positive alignment of value to such prosocial norms is associated with lower future stock returns. Heightened E&S concerns in prosocial norms reduce the level of corporate investments and weaken the link between corporate investments and Tobin's $q$ in the short-run, but ultimately improve both the level and efficiency of corporate investments in the very long-run.

***Keywords:*** *Social Norms, Prosocial Norms, ESG, Natural Language Processing (NLP), Large Language Models (LLM), Word Embeddings, Textual Analysis, Financial History*
***JEL Classification:*** *C82, E44, G12, G31, G32, M14, N00*

*Joel Houston (joel.houston@warrington.ufl.edu), Sehoon Kim (sehoon.kim@warrington.ufl.edu), and Boyuan Li (boyuan.li@warrington.ufl.edu) are with the University of Florida, Warrington College of Business, Department of Finance, Insurance, and Real Estate. We thank Tuugi Chuluun, Veljko Fotak, Adam Kolasinski, Nitish Kumar, Hao Liang, Alejandro Lopez-Lira, Pedro Matos, Nora Pankratz, Yuehua Tang, Aaron Yoon, Xintong Zhan, and conference/seminar participants at the 2025 American Finance Association Conference, the 2024 Financial Management Association Conference, the 2024 Financial Intermediation Research Society (FIRS) Conference, the 2024 UN PRI Academic Network Conference, the 2024 China International Conference in Finance (CICF), the 2024 Global Research Alliance for Sustainable Finance and Investment (GRASFI) Conference, the 2024 Conference for Financial Economics and Accounting, the 2024 Loyola Financial Ethics Conference, the 2024 NTHU-UNSW Symposium on Sustainable Finance and Economics, the 2023 Charleston Conference, and the University of Florida. We also thank Clarivate/ProQuest for data and computational resources. This paper was previously circulated under the title "*One Hundred and Thirty Years of Corporate Responsibility*."

# 1. Introduction

Social norms are critical in influencing economic activity by dictating the societal nexus of contracts. In the literature, such norms are often assumed to vary across countries but fixed over time (see, e.g., Liang and Renneboog, 2017). However, social norms also arise and evolve in the course of history along with socioeconomic and political developments. In recent years, for instance, environmental, social, and governance (ESG) concerns and norms around corporate social responsibility, or "prosocial norms," have risen in prominence among corporations and investors (see Hartzmark and Sussman, 2019; Krueger, Sautner, and Starks, 2020; Bolton and Kacperczyk, 2021, 2022; Pastor, Stambaugh, and Taylor, 2022). However, the difficulty of measuring the multi-faceted contexts of social norms makes it challenging to understand their determinants, the importance of alignment with such norms, or how they affect asset prices and corporate policies.

In this paper, we study the evolution of prosocial norms and their financial and real implications. We overcome the challenges of measuring prosocial norms by employing advanced natural language processing (NLP) techniques that account for the evolution of environmental and social (E&S) context in 130 years of historical news text. Specifically, we train word embedding language models on different time periods throughout the corpus sample and use them to span extensive and time-varying E&S dictionaries. Applying these dictionaries to millions of business/economics news articles from The Wall Street Journal and New York Times, we construct a novel text-based time-series Prosocial Norm Index (PNI) that measures public attention and social norms about E&S issues around businesses.

The time-series of PNI and its E&S components, PNI-E and PNI-S, reflect different aspects of environmental and social concerns that were prevalent at different points in history. Using time-varying language models, our methodology is able to construct these measures without relying on the econometrician's subjective judgment regarding which topics or issues should be reflected in the index during different periods. The index also enables us to derive firm-level measures of "alignment with prosocial norms," where we compute stock-month level PNI exposures from monthly stock returns on a rolling-window basis. Motivated by recent studies that show that shocks to green preferences and the demand for green assets explain the return premium of green over brown assets (see Pastor, Stambaugh,

and Taylor, 2021; Pedersen, Fitzgibbons, and Pomorski, 2021; Pastor et al., 2022), we use firm-level PNI exposures as demand-based proxies for *aligned values*. These measures are free of arbitrary selection of topics that have led to widespread ambiguity and disagreement across contemporary ESG ratings (see Berg, Fabisik, and Sautner, 2021; Berg, Kölbel, and Rigobon, 2022; Berg, Kölbel, Pavlova, and Rigobon, 2023).

We first explore the basic properties of PNI and validate our measure by inspecting its relationships with several macroeconomic and socio-political variables (see Jordà, Schularick, and Taylor, 2016; Piketty, Saez, and Zucman, 2018; Poole and Rosenthal, 1984; Azzimonti, 2018, and the National Bureau of Economic Research (NBER)). Going from periods of macroeconomic and social instability to times of relative prosperity, we observe a migration of public attention and concern from social to environmental issues. We find that public attention to social issues arise amid internal economic and social instability, signified by high levels of PNI-S during times of high volatility in real GDP growth, high unemployment, frequent recessions, and high wealth inequality. In contrast, attention to environmental issues (i.e., high levels of PNI-E) is afforded by society during times of relative stability and prosperity, but positively correlated with partisanship and political polarization. Further validating our measure, we find in recent years that the PNI-E component closely tracks a news-based measure of climate policy uncertainty recently developed by Gavriilidis (2021).

Next, we use PNI to answer first-order questions about the financial and real economic consequences of society's shift toward and away from prosocial norms. Specifically, we examine the cross section of stock returns and corporate investment in relation to PNI and firm exposure to PNI. In our analysis of stock returns, we utilize PNI and monthly stock returns from CRSP dating back to the mid-1920s, highlighting the suitability of our methodology to long-sample studies of social norms. Specifically, we measure each stock's beta with respect to monthly innovations in PNI on a five-year rolling-window basis. This measure of "PNI exposure" serves as a demand-based proxy for a firm's "alignment with prosocial norms," which is rooted in the assumption that a firm is likely to be more aligned with prosocial norms if its stock co-moves positively with such norms. On average, over our extensive sample period, we find no evidence that stocks that are highly and positively exposed to PNI earn higher subsequent returns than stocks that are less or negatively exposed. If anything, our results suggest that stocks with positive exposure perform worse when the level of PNI is low.

We also examine the impact of prosocial norms on the level and efficiency of corporate investment. Both reduced-form regressions as well as impulse responses from panel vector auto-regressive (VAR) models suggest that heightened public concerns about E&S issues dampen business activity in the short-run, indicated by lower capital expenditures in the year immediately following an increase in PNI. In particular, we find that elevated E&S concerns are negatively associated with corporate investment efficiency in the near future, implied by weakening links between corporate investment and Tobin's $q$. However, the VAR impulse response functions also enable us to delineate longer-term effects beyond 1-2 years in the immediate future. The long end of the impulse responses suggest that both the level and efficiency of corporate investments not only recover but eventually improve over a ten-year horizon. These findings underscore and contrast the short-term and long-term consequences of social norm-alignment. The combination and contrast of the stock return analysis and long-run real effects on corporate investments suggest that markets may not be able to quickly and fully price the long-term implications of prosocial norm-alignment.

By exploiting a rich historical time-series, our findings contributes to the growing literature on the financial market implications of corporate responsibility and ESG performance. Given the rise of ESG and climate finance as a recent phenomenon, studies that examine investor ESG demand and its relationship with economic conditions have relied on short sample periods around specific events or shocks (see Hartzmark and Sussman, 2019; Döttling and Kim, 2024). Cross-sectional asset pricing studies have also focused on recent sample periods due to the availability of data on environmental performance (see Bolton and Kacperczyk, 2021, 2022; Hsu, Li, and Tsou, 2023; Zhang, 2023). In particular, the recency of the literature poses a fundamental challenge in assessing the long-term consequences of ESG or responsible corporate practices. This is an important problem given that one of the key premises for ESG propositions is that it creates long-term value (see Edmans, 2023). Our study contributes to overcoming this challenge in the literature and to generalizing this line of research to broader prosocial norms and longer sample periods.

Our work also complements recent studies that analyze granular text data to shed light on how companies incorporate stakeholder values into their policies. For example, Rajan, Ramella, and Zingales (2023) analyze corporate letters to shareholders and document that firms have recently been proclaiming environmental and social goals in their purported

corporate objectives. Li, Shan, Tang, and Yao (2023) and Sautner, van Lent, Vilkov, and Zhang (2023a,b) apply climate-related dictionaries to earnings call transcripts and construct firm-level measures of climate risk exposures. Rouen, Sachdeva, and Yoon (2023) analyze the evolution of content in corporate ESG reports over the past decade. Our study contributes to this literature along two important dimensions. First, by going back to the late 19th century (or early 20th century for firm-level analysis), we dramatically extend the time-series coverage beyond what the literature has been able to examine. Second, by applying NLP techniques based on *word embeddings* to different time periods, our PNI measure incorporates a wide variety of contexts related to E&S issues, accommodating the historical evolution and various aspects of prosocial norms.

Finally, our study contributes to the broader literature that applies NLP techniques in financial economics. Earlier studies in this area use "bag-of-words" approaches that utilize word count frequencies to measure aggregate uncertainty or firm-level risk exposures (see Baker, Bloom, and Davis, 2016; Loughran and McDonald, 2016; Manela and Moreira, 2017; Hassan, Hollander, Lent, and Tahoun, 2019). With the proliferation of NLP techniques and big text data, the literature has also adopted a variety of topic modeling and prediction-based word embedding techniques to understand rich contexts related to economic sentiment, business cycles, or geopolitical risk (see Binsbergen, Bryzgalova, Mukhopadhyay, and Sharma, 2023; Bybee, Kelly, Manela, and Xiu, 2023; Hirshleifer, Mai, and Pukthuanthong, 2023a,b; Jha, Liu, and Manela, 2023). Recent studies also explore the implications of large language models (LLMs) based on generative AI (e.g., ChatGPT) for asset prices and monetary policy (see Lopez-Lira and Tang, 2023; Hansen and Kazinnik, 2023; Jha, Qian, Weber, and Yang, 2023). Our work adds to this literature by employing a novel time-varying word embedding approach recently adopted in the literature (see Bandyopadhyay, Mai, and Pukthuanthong, 2023a,b), which allows the econometrician to model the historical evolution of context free of subjective judgment.

The paper is organized as follows. In Section 2, we motivate our study by briefly describing the evolution of prosocial norms in the United States over the past 130 years. In Section 3, we discuss our data and methodology used to construct PNI. In Section 4, we explore and validate the proprties of PNI, and use it to examine the financial implications

of value alignment with social norms and the real effects of shifting prosocial norms. In Section 5, we summarize our findings and conclude.

## 2. Prosocial Norms Throughout U.S. History

Despite recent interest and debates around ESG and corporate sustainability, the general concern that corporations should have a broader responsibility to the general public beyond maximizing profits and shareholder wealth is not a new phenomenon. Throughout U.S. history, the call for corporate responsibility and prosocial firm behavior has surfaced repeatedly and continuously since at least the end of the Gilded Age in the 1890s.

For example, in 1889, Andrew Carnegie wrote in the "Gospel of Wealth" about the responsibility of the newly minted self-made upper-class to reduce wealth inequality. In 1931, Adolf Berle and Merrick Dodd debated the role of the corporation in serving society and public interest, and the place of shareholders in relation to labor, customers, and the community. This series of exchanges became well-known as the Berle-Dodd debate. Such public discourse around the externalities of business and the firm's social responsibilities continued through 1970 when Milton Friedman famously emphasized that the social responsibility of business was to increase its profits, which from the 1980s to 2000s became widespread as the "Friedman Doctrine" of shareholder supremacy. In 2004, a partnership between UN Global Impact and Switzerland launched the "Who Cares Wins" initiative, in which the now-widespread acronym, Environmental, Social, and Governance (ESG), was coined as a movement to align sustainability goals and financial value. Figure 1 illustrates the historical timeline of these anecdotal examples that highlight the recurrence of public interest in corporate responsibility and the evolution of prosocial norms in the U.S.

[Insert Figure 1 here]

The context in which corporations have been called on to serve the interest of the general public has also been multifaceted throughout history, ranging from social issues related to epidemics/pandemics (e.g., Spanish flu, COVID-19), the social fallout of wars (e.g., World Wars I and II), poverty amid economic crises (e.g., the Great Depression), civil

rights movements, to environmental externalities such as industrial water and air pollution or climate change. This rich historical backdrop and the complex, evolving nature of environmental and social problems call for a broader, longer-term perspective and a generalized framework of understanding the role of prosocial norms, which contrasts with what the previous literature has only been able to offer in smaller samples and shorter time periods.

Undoubtedly, building such a framework is methodologically and empirically challenging due to data availability and consistency, as well as selection biases that would contaminate any attempt to manually choose and construct new data from historical sources. However, many of these challenges can by overcome by coupling rich text data whose availability spans longer sample periods than what most standard finance and sustainability databases can provide, with state-of-the-art natural language processing (NLP) techniques to parse and analyze such nuanced textual information. In the following section, we detail our data and methodology used to achieve this goal.

# 3. Data and Methodology

## 3.1. Data Sources

Our primary data source is an extensive text corpus of historical news articles from The Wall Street Journal (WSJ) and The New York Times (NYT) spanning 130 years back to 1890. We obtain digitized copies of these articles from ProQuest TDM Studio. Our initial pre-processing of the raw articles involves removing punctuations, digits, and special characters. We then lemmatize each word by transforming it to its base form based on the word's *part-of-speech*.[1] We further remove all stop words that are frequently used but do not carry substantial meaning by themselves.[2] We also require that within each article, at least 1% of words are related to the seed words, "*business*," "*economy*," or "*corporation*," based on a dictionary generated from *word embeddings* as we describe in more detail below. Lastly, we exclude all articles with fewer than 100 content words, following Hirshleifer

[1]For example, "*faster*" and "*cars*" are transformed into "*fast*" and "*car*," respectively.

[2]To identify stop words, we use the "Expanded stop words list" developed by Matthew L. Jockers (https://www.matthewjockers.net/macroanalysisbook/expanded-stopwords-list/).

et al. (2023a). After pre-processing, we are left with approximately 4 million news articles related to business and the economy.

We inspect several historical macroeconomic variables in our time-series analysis of PNI. We obtain real GDP growth rates and unemployment rates from the Jordà et al. (2016) Macrohistory database, and National Bureau of Economic Research (NBER) recession indicators from the St. Louis Fed. We also obtain wealth inequality, measured as the share of total household wealth held by the top 1% percentile of households, from Piketty et al. (2018). We use two measures of political frictions: congressional polarization (see Poole and Rosenthal, 1984, and Voteview.com) and news-based partisan conflict (see Azzimonti, 2018). We adopt two additional news-based indices: climate policy uncertainty from Gavriilidis (2021), and geopolitical risk from Caldara and Iacoviello (2022).

Finally, stock returns (from 1926 and onward) and firm fundamentals (from 1960 and onward) are obtained from CRSP/Compustat.

## 3.2. Methodology

The objective of this paper is to construct a textually derived measure that not only quantifies public attention and social norms about environmental and social (E&S) issues around businesses, but also incorporates the historical evolution of the context of prosocial norms. To achieve the latter, it is critically important that the researcher's subjective bias in terms of salience or familiarity regarding historical events does not confound the measurement. For this reason, we carefully design an NLP methodology that minimizes the need to manually select topics from the historical corpus.

Our methodology for constructing our Prosocial Norms Index (PNI) follows three stages. The first stage is to generate time-varying language models that can account for the evolution of linguistic context over time. This step is critical for our study as we aim to create an index capturing discourse in issues for which the contexts are in flux and multifaceted. The second stage is to use these language models to generate a time-varying set of dictionaries without infusing our subjective judgment of how those dictionaries ought to change over time. The third stage entails constructing PNI using these dictionaries.

### 3.2.1. Time-Varying Language Models

To implement the first stage of our methodology, we use the skip-gram implementation of Word2Vec, a widely used NLP technique that relies on prediction-based word embeddings. This algorithm produces vector representations of words and phrases in a corpus by using the co-occurrence of words to predict context words surrounding a target word. Specifically, it solves for an objective function that maximizes the probability that any "context" word in a corpus appears within some radius of a given target word. The byproduct of this optimization is the full vector space of all words and phrases in the corpus, where words with similar contextual meanings have similar vector representations. The distances between these vectors, or "cosine similarities," can be used to measure how similar words are with one another. The key strength of this algorithm compared to traditional "bag-of-words" approaches is that it captures the context in which a word or phrase is used, such that it can identify contextually similar phrases even if they are not used together in proximity.

This algorithm is well-suited for the purpose of our study. Unlike recent studies that rely on pre-trained language models such as BERT or ChatGPT (see Jha et al., 2023; Lopez-Lira and Tang, 2023; Hansen and Kazinnik, 2023), we are interested in the evolution of context over time. Therefore, we train separate Word2Vec models for each decade throughout the sample period (e.g., 1890–1899, 1900–1909, ... , etc.) to generate our own set of time-varying language models. This allows us to capture the historical evolution of context in which the public has been concerned about environmental or social issues.

When training our models, we identify bi-grams and tri-grams in our corpus to ensure that the Word2Vec algorithm accounts for contexts in which a word is used as part of a multi-word phrase.[3] As hyperparamaters needed for training the language models, we use a vector dimension size of 300 and the standard context radius (i.e., window length) of 10.[4] In addition, we drop words that appear in the corpus less than five times and run the skip-gram algorithm 20 times (instead of 5 times, as in default) to enhance our models' performance.

[3]For example, "*new*," "*york*," "*city*" are three separate words, but are frequently used together in the three-gram, "*new_york_city*."

[4]The model output is robust to alternative window lengths of 5 and 15.

### 3.2.2. Subjectivity-Free, Time-Varying Dictionaries

Next, we use these decade-specific language models to generate time-varying dictionaries that contain keywords that are closely related to E&S issues, as semantically used in each decade. To ensure that the time-variation in the dictionaries is not driven by our subjective judgment of which environmental and social problems ought to have been important aspects of social norms in different time periods, we start from a "universal," or *time-invariant* set of seed words that we can feed to the language models. We choose these initial seed words carefully, such that they not only appear in every decade of our corpus, but can also be unambiguously interpreted as issues pertaining to E&S issues. These seed words are *pollution*, *inequality*, and *discrimination*. These words have been used throughout history with relatively little semantic ambiguity and are categorically general enough such that they may be related to many facets of the same broad category. For example, the word "*pollution*" may be related to many types of environmental externalities such as water pollution, air pollution, or greenhouse gas emissions. Similarly, "*inequality*" or "*discrimination*" may be related to class conflicts along various dimensions such as wealth, gender, or race. Depending on which specific topics were particularly relevant in different points of history, these same seed words would span different sets of contextually similar words when they are fed to decade-specific language models.

We use the language models to span time-varying dictionaries by identifying the top 100 words and phrases with the highest cosine similarity scores for each seed word within each decade. To ensure that our dictionaries are not polluted by generic words or phrases that are used often in sentences, we drop keywords that are in the top 50 in terms of frequency of use, but also in the bottom 50 in terms of their cosine similarity scores. We use this automated approach to trim the dictionary to minimize the influence of subjective judgment on dictionary context.

Figure 2 illustrates how the context of discourse evolves over time by showing word clouds for select years. The word clouds visualize the top 500 keywords that are most closely related to each of our seed words in terms of their cosine similarity scores, weighted by their frequency of occurrence in the corpus. For select years, the figure shows word clouds of keywords generated from the seed word "*pollution*" in Panel A, and from the seed words "*discrimination*" and "*inequality*" in Panel B.

[Insert Figure 2 here]

The word clouds clearly illustrate that our methodology captures and allows for an ongoing evolution in the context of each seed word. For example, pollution-related discourse in business/economy-related news has evolved from focusing on issues related to water pollution in 1920, to increasingly discussing air pollution in 1970, to heavily discussing climate change and greenhouse gas emissions in 2020.

#### 3.2.3. Construction of the Prosocial Norms Index (PNI)

Finally, we use the time-varying dictionaries to construct the Prosocial Norms Index, or PNI. For each news article, we first count the total number of occurrences of the keywords contained in the dictionary, and divide it by the total word count of the article.[5] PNI is then computed as the monthly average of this ratio across all news articles in that month. For subsequent analysis, we also construct environmental and social sub-indices, PNI-E and PNI-S, based on dictionaries spanned separately from the environmental seed word "*pollution*" or social seed words "*inequality*" and "*discrimination*."

Figure 3 illustrates and summarizes the methodology for constructing our monthly PNI index that we have discussed in detail above.

[Insert Figure 3 here]

# 4. Results

## 4.1. Properties of PNI

We begin by exploring the properties of PNI. Figure 4 presents the full monthly time-series of PNI from January 1889 to March 2023. Panel A shows the overall PNI measure, and Panel B shows PNI-E and PNI-S separately.

[5]To account for the fact that we use *one* environmental seed word, "*pollution*," and *two* social seed words, "*inequality*" and "*discrimination*," we average the word counts of keywords spanned from the two social seed words before aggregating the environmental and social word counts.

[Insert Figure 4 here]

Several notable patterns are observed. Over the past 130 years, there have been several waves of high-PNI periods. In the first half of the 20th century, long before formalized concepts of corporate social responsibility (CSR) or ESG, our PNI measure indicates that there had already been much public discourse about environmental and social issues shaping prosocial norms. Separate plots of PNI-E and PNI-S reveal that much of this early concern was driven by social issues related to inequality and discrimination. This period, which includes the end of World War I, the Spanish Flu, and the 1929–1939 great depression, was indeed marked by poverty, rampant inequality, and labor rights movements. In the post-war era after 1950, PNI-S subsides – with a spike around the Civil Rights Movement in the 1960s – while PNI-E begins to dominate the index. In particular, PNI-E captures conversations in the mid-1960s about protecting the environment, sharply spiking after the establishment of the U.S. Environmental Protection Agency (EPA) in late 1970. This was also the period when widespread debates on corporate social responsibility began to take hold, marked by the "Friedman (1970) doctrine." During the recent three decades since 1990, there has been a gradual uptick in PNI, PNI-E, and PNI-S as ESG, climate change, gender and racial diversity came front and center in public discourse among investors and business leaders alike. This recent period was punctuated by verifiable spikes around the Clean Air Act Amendments in 1990, the Kyoto Protocol in 1997, the Copenhagen Climate Change Conference in 2009, the Paris Agreement in 2015, the Black Lives Matter (BLM) movement in 2020, as well as COP26 and the U.S. Infrastructure Bill in 2021.

[Insert Table 1 here]

Table 1 describes the time-series properties of PNI and its components. The mean PNI value throughout the entire sample period is 0.05, indicating that 0.05% of the average article's text consists of keywords that belong to our decade-specific E&S dictionary. The components, PNI-E and PNI-S, have correspondingly lower means of 0.02 and 0.03, respectively. The median values are similar to the means, indicating that the indices are not significantly skewed, despite their occasional spikes. The standard deviation is 0.02 for the overall index as well as its components, implying that our methodology generates consistent

variability across the topics it covers. The minimum and maximum index values are also reported separately for two sub-periods: pre-1970 and post-1970. All of the indices exhibit wide ranges in both periods, consistent with long-run waves of prosocial norms driven by public discourse related to environmental and social issues surrounding businesses.

## 4.2. PNI and Macroeconomic Conditions

To systematically understand the forces that drive fluctuations in prosocial norms, we start by examining the time-series relationships between our indices and several macroeconomic variables. Figure 5 illustrates how PNI-E and PNI-S are associated with real GDP growth (Panel A), unemployment (Panel B), recessions (Panel C), and wealth inequality (Panel D). Columns 1 to 4 of Table 2 present formal time-series regressions of these associations (see Equation 1).

$$PNI_t = \alpha + \beta \cdot Macro\ variable_t + \epsilon_t \tag{1}$$

Figure 5 paints a clear picture that helps explain the historical patterns in PNI-E and PNI-S. During periods of internal economic strife indicated by high volatility in real GDP growth, high unemployment, frequent recessions, and high wealth inequality, public discourse focuses more on social externalities such as inequality and discrimination (i.e., high PNI-S), rather than on environmental problems (i.e., low PNI-E). In particular, PNI-S very closely tracks wealth inequality measured as the top 1%'s share of total household wealth (see Piketty et al., 2018). In contrast, the stabilization of the economy and society is followed by a shift in public focus to environmental issues, elevating PNI-E to higher levels.

[Insert Figure 5 here]

Table 2 shows that these relations are statistically and economically significant. Columns 2 to 4 show that PNI-E is negatively and significantly associated with unemployment, recession period dummies, and wealth inequality, whereas PNI-S is positively and significantly associated with all of these variables. Column 1 shows that PNI-E is negatively associated with real GDP growth, indicating that public concerns regarding the environment rises in a maturing economy with slowing growth.

[Insert Table 2 here]

Figure 6 and Table 2 (Columns 5 to 9) further examine the relationships between PNI and several sources of political frictions. These results show that political polarization based on congressional votes (both House and Senate) and the news-based partisan conflict index by Azzimonti (2018) are both positively and significantly correlated with PNI-E. On the other hand, their relationships with PNI-S are mixed, exhibiting negative associations during the pre-1970 sample but positive associations post-1970. Corroborating the validity of PNI, our indices closely track a news-based climate policy uncertainty index by Gavriilidis (2021). This is consistent with the idea that the recently positive associations between the PNI-E/S and political frictions may be due to rising contention over environmental and social policy issues.

[Insert Figure 6 here]

Unlike internal political conflicts, external geopolitical risks are not as closely linked to PNI. Although Column 9 of Table 2 shows that PNI-E (PNI-S) is negatively (positively) associated with the news-based geopolitical risk measure by Caldara and Iacoviello (2022), visual inspection of Figure 6 (Panel D) suggests that these statistical associations are mainly driven by World Wars I and II.

## 4.3. Value of Aligned Values: PNI Exposure and Stock Returns

Armed with a macro-founded validation of PNI, we next apply our measure to examine whether alignment with prosocial norms affects asset prices as a first-order question. We do this by measuring firms' exposures to PNI, and testing whether firms with higher or lower PNI exposure earn higher or lower stock returns.

An underlying assumption of this exercise is that a firm is likely to be more aligned with prosocial norms if its stock co-moves positively (negatively) with public attention and social norms about E&S issues around businesses. This is motivated by recent studies that theoretically and empirically show that shocks to green preferences – and thus the demand for green assets – explain the return premium of green over brown investments (see Pastor

et al., 2021, 2022; Pedersen et al., 2021). If innovations in PNI can be interpreted as common shocks to prosocial preferences, we can interpret their correlation with a stock's returns (i.e., PNI exposure) as a demand-based proxy for the firm's alignment with prosocial norms. PNI exposure not only provides unprecedented time-series coverage, but is also free of arbitrary or manual selection of specific aspects of E&S issues that often result in ambiguity and disagreement across third-party ESG ratings (see Berg et al., 2021, 2022, 2023).

We measure stock-month level PNI exposure by estimating a five-year rolling window regression of the stock's return in excess of the risk-free rate, on the Fama and French (1993) three-factor model augmented with monthly innovations in PNI as an additional factor. From these regressions, we obtain the time-varying beta coefficients on the PNI innovation term, or PNI-betas, as our measure of PNI exposure.

$$r_{i,t} - rf_t = \alpha_i + \beta_i \cdot \Delta PNI_t + \gamma_i \cdot [rm_t - rf_t] + \delta_i \cdot SMB_t + \eta_i \cdot HML_t + \epsilon_{i,t} \qquad (2)$$

We start by inspecting PNI exposures in different industries over different time periods. As the context of prosocial norms changes over time, so will the universe of firms that are most positively or negatively impacted by changes in such norms. We summarize this evolution by tabulating the Fama-French 30 industries that consisted of the most firms in the highest and lowest PNI exposure decile groups over the past five two-decade periods (i.e., 1930–1949, 1950–1969, ... , 2010–present).[6]

[Insert Table 3 here]

The summary is reported in Table 3, which illustrates how the set of industries affected by public concerns about environmental and social issues have shifted over time. Panels A to C each show the top and bottom three industries in terms of exposures to PNI, PNI-E, and PNI-S, respectively. Panel B, for example, demonstrates that fossil fuel and other greenhouse gas emitting industries have increasingly become negatively exposed to PNI-E (i.e., their stocks perform poorly when the public is more concerned about environmental issues). Panel C also shows that industries that were at the center of labor disputes and protests in the late 19th

[6]The sample period is restricted to after 1930 due to the availability of stock returns from CRSP.

and early 20th century, such as textiles and steel works, were among the industries most negatively exposed to PNI-S during those historical periods. Overall, the firm-level exposures derived from our indices produce sensible patterns that successfully capture historical context.

We now examine whether stocks with greater PNI exposure command higher or lower future returns. We run regressions of monthly stock returns on PNI exposures, obtained from five-year rolling regressions as detailed above (see Equation 2). PNI exposure estimated over the five-year window ending month $t-1$ is used to explain returns in month $t$. The regressions are run on the full sample period from February 1931 to January 2023, and also on subsamples covering the pre-1970 and post-1970 periods when concerns of social and environmental issues each dominated the discourse driving prosocial norms, respectively. We include a vector of standard control variables, denoted as $\mathbf{X}$, including lagged values of firm size (i.e., market capitalization), lagged monthly returns, past twelve months' returns skipping a month (i.e., momentum), past three year's monthly returns skipping a year (i.e., reversal), idiosyncratic volatility computed from daily residuals from the Fama and French (1993) three-factor model estimated over the previous twelve months, and the average bid-ask spread.[7] We include stock and month fixed effects, and cluster standard errors at the stock level.

$$r_{i,t} = \beta \cdot PNI\ exposure_{i,t-1} + \gamma' \cdot \mathbf{X}_{i,t-1} + \eta_i + \theta_t + \epsilon_{i,t} \tag{3}$$

[Insert Table 4 here]

The results are reported in Table 4. Panel A uses the raw continuous exposures as explanatory variables, while Panel B uses dummy variables indicating whether exposure is positive. Throughout all sample periods, we find no evidence that positive exposure to the overall PNI measure is associated with higher future stock returns. If anything, we find that positive exposure hurts stock performance when the index level is low. For example, in the pre-1970 period when the level of PNI-E is low, stocks with more positive exposure to PNI-E tend to perform worse. Similarly, in the post-1970 period when PNI-S is low, stocks more positively exposed to PNI-S tend to under-perform. Moreover, this under-performance

[7] In untabulated analysis, we also add book-to-market, profitability, asset growth, and accruals, which restricts our sample period to the post-1960 period due to the availability of fundamentals data. The coefficients we obtain for the PNI exposures remain identical to those from our baseline post-1970 regression.

is not compensated for during times of high index levels: Stocks with positive exposures to PNI-E (PNI-S) perform no better than other stocks even when the index level is high during the (pre-) post-1970 period. These results suggest that it costs more than it pays to invest in stocks with positive PNI-betas: These stocks do poorly when public discourse does not focus on environmental or social issues, and they do not do particularly well even when public concerns on these issues are elevated.

[Insert Figure 7 here]

This result is also encapsulated in Figure 7, which plots the cumulative return on a portfolio strategy that, each month, longs the top decile of stocks and shorts the bottom decile of stocks sorted on their previous month's PNI exposures. The figure illustrates detailed time-series variation in the profitability of such a "norm-aligned" investing strategy. Consistent with the regression results in Table 4, Panel A of Figure 7 shows that the strategy does not deliver positive returns over the very long-run (i.e., entire sample period) nor over the pre-1970 or post-1970 sub-periods. Panel B shows that a PNI-E beta long-short strategy has negative cumulative returns in the pre-1970 era, while also failing to generate positive cumulative returns in the post-1970 era. The latter is due to the fact that even long periods of good PNI-E beta performance (e.g., decade-long periods in the 1980s or 2000s) are followed by sharp and large reversals (e.g., early 1990s or early 2010s). Panel B also shows that a PNI-S beta long-short strategy delivers flat returns before 1970 and performs negatively after 1970. In short, aligning with prosocial norms does not necessarily deliver long-term value and is therefore a myopic strategy for maximizing shareholder value.

Analogizing alignment with prosocial norms to corporate responsibility, one might view the evidence in the literature showing a positive association between corporate responsibility and stock returns to be limited to a few sample periods throughout history and hard to generalize. A first-pass application of our PNI methodology on the cross-section of stock returns casts doubt on the often-debated idea that "companies can do well by doing good." However, given that these results are based on the relationships between stock returns and PNI exposures with immediate lags, it is possible that there are positive, *real* changes to firms associated with prosocial norms that occur over long horizons, which the market merely fails to price. Moreover, econometrically estimating long-term value

implications based on equilibrium asset prices is notoriously difficult. To shed more light on the *long-term* consequences of prosocial norms using our PNI measure, we next turn to examining the real effects of innovations in PNI.

## 4.4. PNI and Corporate Investments

In addition to implications for asset prices, an equally important first-order question is whether prosocial norms have real effects on the decisions of firms. As one of the most important decisions made by firms, we examine both the level and efficiency of corporate investments. We study these real effects both in the short-run and long-run by complementing reduced form regressions with panel vector auto-regressive (VAR) models. These analyses over different time horizons are also informative about the ability of financial markets to price long-term changes in investments.

On one hand, heightened prosocial norms may require firms to invest more in creating socially equitable organizations or implementing environmentally friendly practices, to satisfy investors and customers who exhibit strong prosocial preferences or to comply with changing regulatory requirements. However, these areas may not necessarily be where the best growth opportunities reside in the present day, meaning that such changes in investments may be inefficient. Greater concerns about environmental and social externalities created by businesses may also curtail economic activity due to regulatory uncertainty and therefore suppress investment in the short-run (see Bartram, Hou, and Kim, 2022).

On the other hand, a greater societal emphasis on corporate responsibility may push firms to implement E&S friendly policies that lower their risk profiles and thus their costs of capital. These policies may also enable them to attract a more productive workforce in the long-run. As such, firms may be able to recuperate the short-term compromises they make for conducting business responsibly through longer-term efficiency gains. Indeed, such sources of long-term value often underlie the advocacy for corporate responsibility. Yet, short- vs. long-horizon delineations are difficult and largely absent in the recent literature on ESG and sustainable finance due to lack of long-sample data. We leverage the extensive sample period for which we are able to construct our PNI measure to help draw novel insights on this issue.

#### 4.4.1. OLS Regressions

We first examine whether higher levels of PNI are associated with greater or smaller corporate investments in the short-run. To do this, we run regressions of corporate investments, measured as capital expenditures scaled by lagged assets, on lagged values of PNI and/or its year-to-year innovations (see Equation 4). Alternatively, we regress investments on the components of PNI, namely PNI-E or PNI-S and/or their innovations. We control for a vector of firm controls, denoted **X**, including Tobin's *q*, return on assets (ROA), long-term debt as a fraction of assets, and firm size measured as the log of total assets. We also control for firm and industry-by-*decade* fixed effects. Standard errors are clustered at the firm level.

$$\frac{Capex_{i,t}}{Assets_{i,t-1}} = \beta_1 \cdot PNI_{t-1} + \beta_2 \cdot \Delta PNI_{t-1} + \gamma' \cdot \mathbf{X}_{i,t-1} + \eta_i + \theta_{j,[t-10 \to t-1]} + \epsilon_{i,t} \quad (4)$$

The results are reported in Panel A of Table 5. The key takeaway is that higher levels of PNI are associated with significantly lower investments in the following year, and that this is primarily driven by PNI-E rather than by PNI-S. These indicate that heightened environmental concerns are particularly not conducive to business activity. This result may be explained by the likelihood that broad concerns about corporate environmental profiles are often associated with heightened uncertainty about future environmental regulations. These regulations often require firms to make large capital investments for compliance and bear significant clean-up costs upon violations, more so than social norms or rules regarding representation or equality. From a real-options perspective, it is intuitive that firms may reduce or postpone investments in the face of greater uncertainty about such regulations. Controlling for Tobin's *q* and industry-by-decade fixed effects mitigate concerns that the reduction in investments may be driven by worsening investment opportunities, or by the long-run transition from the pre-war to post-war era that coincided with a structural shift in economic stability and prosperity.

[Insert Table 5 here]

Next, we further examine whether increases in PNI are associated with changes in investment efficiency. To test this, we augment the investment regressions by further interacting

PNI with lagged Tobin's $q$, while retaining Tobin's $q$ as a standalone control variable. In these tests, we include more granular fixed effects at the firm and industry-by-*year* levels, the latter subsuming the time-series PNI variable.

$$\begin{aligned}\frac{Capex_{i,t}}{Assets_{i,t-1}} = \beta_1 \cdot PNI_{t-1} \times Tobin's\ q_{i,t-1} + \beta_2 \cdot \Delta PNI_{t-1} \times Tobin's\ q_{i,t-1} \\ + \gamma' \cdot \mathbf{X}_{i,t-1} + \eta_i + \theta_{j,t} + \epsilon_{i,t}\end{aligned} \tag{5}$$

Panel B of Table 5 reports the results. Higher levels of PNI and greater increases in PNI are both associated with a decrease in investment efficiency. The coefficients on the interaction terms are negative and significant throughout most specifications. Together with the fact that Tobin's $q$ is positively and significantly associated with corporate investments, these results indicate that higher levels and greater innovations of PNI substantially weakens the positive link between corporate investments and Tobin's $q$. The one exception is the interaction term between the *level* of PNI-E and $q$, which has a positive coefficient. This contrasts with the negative coefficient on the interaction term between the *change* in PNI-E and $q$. One explanation for this could be that the *level* of PNI-E reveals information about which investment opportunities are compliant with new regulatory frameworks, whereas *changes* in PNI-E reflect regulatory uncertainty and confusion about the viability of growth options elsewhere. Therefore, while greater *changes* in PNI-E make investments more inefficient, higher *levels* of PNI-E may result in smaller but more targeted investments.

Overall, these results are consistent with a negative impact of prosocial norms on corporate investment and investment efficiency in the short-run.

#### 4.4.2. Impulse Response Functions from Panel VARs

To help draw causal inferences about the short-term effects documented above and to delineate longer-term impacts that are difficult to estimate from reduced form OLS regressions, we estimate panel vector autoregressions (VARs) and plot orthogonalized impulse response functions. We include corporate investments (i.e., capital expenditures scaled by lagged assets), yearly average PNI, Tobin's $q$, return on assets (ROA), long-term debt over assets (Debt), and log of total assets (Size) as dependent variables, and fit a multivariate panel

regression of each dependent variable on the first lag of itself and on the first lags of all other dependent variables using generalized method of moments (GMM). We control for firm fixed effects by taking first-differences and adjust standard errors for clustering at the firm level.

In Figure 8, we plot the responses of corporate investments with respect to positive unit shocks to our PNI measure (Panel A), and with respect to PNI-E or PNI-S shocks separately (Panel B). Consistent with the evidence in Table 5, the impulse response functions show a sharp reduction in the level of investments in the two years immediately following positive shocks to PNI, amounting to 0.4% as a fraction of assets. The figure also illustrates interesting longer-term dynamics after the initial decline in investments. Investments gradually recover to pre-impulse levels over the next five years, suggesting that shocks to PNI can have persistent effects on corporate investments. However, in the long-run, approximately seven years after the unit shock to PNI, investments fully recover to pre-impulse levels and increase further up to ten years out. Panel B shows that the brunt of the effect arises from shocks to environmental concerns (i.e., PNI-E), where it is clearer that the level of investment initially declines but eventually rises above pre-impulse levels in the long-run.

[Insert Figure 8 here]

To further show the dynamics of investment efficiency with respect to shocks to PNI, Figure 9 plots the investment responses separately for subsamples of firms with high or low ex-ante Tobin's $q$. To allow for the impulse response functions to track average within-firm changes, we classify firms with respect to the cross-sectional median using each firm's first available data point with a valid Tobin's $q$. The left and right charts in Figure 9 each report the investment responses by high and low $q$ firms, respectively. Consistent with the OLS results, the figure shows that the initial negative response by high $q$ firms is larger in magnitude and also more significant than the response by low $q$ firms. This indicates that positive shocks to PNI have a disproportionately negative impact on near-term investments by high $q$ firms, weakening the link between investments and Tobin's $q$ in the short-run. On the other hand, the long-run dynamics in the impulse responses suggest that high $q$ firms eventually invest substantially more than they originally had after ten years, whereas low $q$ firms merely return to their original levels of investment. This indicates that heightened prosocial norms foster an environment in which resources are allocated more efficiently

in the long-run, i.e., firms with better investment opportunities are made to invest more, despite the fact that economic activity and efficiency is initially dampened in the short-run.

[Insert Figure 9 here]

Finally, Figure 10 plots the impulse responses of investments for different subperiods or industry subsamples. In Panel A, the responses are plotted separately for subperiods before and after 2004, the year when the UN coined the term "ESG" for the first time.[8] The plots indicate that corporate investments initially decline in response to positive innovations in PNI during both periods. Interestingly, the long-run, positive dynamics of investments have become more prominent since 2004, after which the concept of ESG investing was popularized. One possible explanation is that since the introduction of ESG investing, prosocial norms have evolved to become more "value-conscious," rather than demanding that companies forgo profitability in the name of corporate responsibility. In Panel B, the responses are plotted separately for firms in the manufacturing / oil & gas / utilities sectors and firms in all other industries. The plots show that the magnitude of the initial negative response of corporate investment is larger among the first group of firms, indicating that pollutive sectors are more susceptible and responsive to stakeholder pressure. After being disrupted, investments by these firms do not fully normalize even in the long-run, suggesting that the immutability of a firm's business model can affect its ability to navigate stakeholder demands.

[Insert Figure 10 here]

# 5. Conclusion

Leveraging state-of-the-art natural language processing (NLP) techniques and applying them to an extensive corpus of historical news text, we construct a historical time-series index dating back to late 19th century that measures public attention and social norms about environmental and social issues around businesses. We call this measure the Prosocial Norms Index (PNI). Our approach incorporates the evolution of linguistic context over time,

[8] See "Who cares wins: Connecting financial markets to a changing world," UN Global Impact (2004).

enabling our measure to reflect different aspects of environmental and social concerns that were prevalent in different points of time. We also derive stock-month level PNI exposures that provide demand-based proxies for firm-level alignment with prosocial norms. Our methodology is inoculated from subjective judgment regarding which topics should be included in the index during different periods, and also free of arbitrary selection of topics that result in widespread ambiguity and disagreement across third-party ESG ratings.

We validate the PNI measure by relating it to macroeconomic variables and inspecting how industries with high or low PNI exposures have shifted throughout history. We also use PNI to study the value implications of norm-alignment, and investigate how changes in prosocial norms impact real firm decisions such as their investments. We find that public attention to social issues arise during times of internal economic and social instability, whereas attention to environmental issues are afforded during times of relative prosperity. At first glance, our analysis relating stock returns to innovations in PNI casts doubt on the notion that corporate responsibility creates value. However, an examination of the impulse responses of corporate investments to positive shocks to PNI suggests that heightened prosocial norms dampen business activity and investment efficiency in the short-run, while eventually improving both in the long-run. Overall, these findings suggest that financial markets may not always fully or correctly price the long-term effects of stakeholder demands on the operations of firms.

Our findings provide researchers with new insights and tools to study long-sample time-series developments in prosocial norms from a historically informed and holistic perspective. We also shed much needed light on our understanding of both the short-term and long-term implications of public discourse and societal demand for corporate responsibility.

# References


Azzimonti, M. (2018). Partisan conflict and private investment. *Journal of Monetary Economics 93*, 114–131.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Bandyopadhyay, A., D. Mai, and K. Pukthuanthong (2023a). AI narrative and stock mispricing. *Working Paper*.

Bandyopadhyay, A., D. Mai, and K. Pukthuanthong (2023b). Diversity narrative and equity in firm leadership. *Working Paper*.

Bartram, S., K. Hou, and S. Kim (2022). Real effects of climate policy: Financial constraints and spillovers. *Journal of Financial Economics 143*(2), 668–696.

Berg, F., K. Fabisik, and Z. Sautner (2021). Is history repeating itself? The (un)predictable past of ESG ratings. *Working Paper*.

Berg, F., J. Kölbel, A. Pavlova, and R. Rigobon (2023). ESG confusion and stock returns: Tackling the problem of noise. *Working Paper*.

Berg, F., J. Kölbel, and R. Rigobon (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance 26*(6), 1315—-1344.

Binsbergen, J. H. v., S. Bryzgalova, M. Mukhopadhyay, and V. Sharma (2023). (Almost) 200 years of news-based economic sentiment. *Working paper*.

Bolton, P. and M. Kacperczyk (2021). Do investors care about carbon risk? *Journal of Financial Economics 142*(2), 517–549.

Bolton, P. and M. Kacperczyk (2022). Global pricing of carbon-transition risk. *The Journal of Finance, forthcoming*.

Bybee, L., B. T. Kelly, A. Manela, and D. Xiu (2023). Business news and business cycles. *The Journal of Finance, forthcoming*.

Caldara, D. and M. Iacoviello (2022). Measuring geopolitical risk. *American Economic Review 112*(4), 1194–1225.

Döttling, R. and S. Kim (2024). Sustainability preferences under stress: Evidence from COVID-19. *Journal of Financial and Quantitative Analysis 59*(2), 435–473.

Edmans, A. (2023). The end of ESG. *Financial Management 52*(1), 3–17.

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

Friedman, M. (1970). A Friedman doctrine: The social responsibility of business is to increase its profits. *The New York Times*.

Gavriilidis, K. (2021). Measuring climate policy uncertainty. *Working paper*.

Hansen, A. L. and S. Kazinnik (2023). Can ChatGPT decipher Fedspeak? *Working paper*.

Hartzmark, S. M. and A. B. Sussman (2019). Do investors value sustainability? A natural experiment examining ranking and fund flows. *The Journal of Finance 74*(6), 2789–2837.

Hassan, T. A., S. Hollander, L. v. Lent, and A. Tahoun (2019). Firm-level political risk: Measurement and effects. *The Quarterly Journal of Economics 134*(4), 2135–2202.

Hirshleifer, D. A., D. Mai, and K. Pukthuanthong (2023a). War discourse and disaster premia: 160 years of evidence from stock and bond markets. *Working paper*.

Hirshleifer, D. A., D. Mai, and K. Pukthuanthong (2023b). War discourse and the cross-section of expected stock returns. *Working paper*.

Hsu, P.-H., K. Li, and C.-Y. Tsou (2023). The pollution premium. *The Journal of Finance 78*(3), 1343–1392.

Jha, M., H. Liu, and A. Manela (2023). Does finance benefit society? A language embedding approach. *Working paper*.

Jha, M., J. Qian, M. Weber, and B. Yang (2023). ChatGPT and corporate policies. *Working paper*.

Jordà, O., M. Schularick, and A. M. Taylor (2016). Macrofinancial history and the new business cycle facts. *NBER Macroeconomics Annual 31*, 213–263.

Krueger, P., Z. Sautner, and L. T. Starks (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies 33*(3), 1067–1111.

Li, Q., H. Shan, Y. Tang, and V. Yao (2023). Corporate climate risk: Measurements and responses. *The Review of Financial Studies, forthcoming*.

Liang, H. and L. Renneboog (2017). On the foundations of corporate social responsibility. *The Journal of Finance 72*(2), 853–910.

Lopez-Lira, A. and Y. Tang (2023). Can ChatGPT forecast stock price movements? Return predictability and large language models. *Working paper*.

Loughran, T. and B. McDonald (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research 54*(4), 1187–1230.

Manela, A. and A. Moreira (2017). News implied volatility and disaster concerns. *Journal of Financial Economics 123*(1), 137–162.

Pastor, L., R. F. Stambaugh, and L. A. Taylor (2021). Sustainable investing in equilibrium. *Journal of Financial Economics 142*(2), 550–571.

Pastor, L., R. F. Stambaugh, and L. A. Taylor (2022). Dissecting green returns. *Journal of Financial Economics 146*(2), 403–424.

Pedersen, L. H., S. Fitzgibbons, and L. Pomorski (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics 142*(2), 572–597.

Piketty, T., E. Saez, and G. Zucman (2018). Distributional national accounts: Methods and estimates for the United States. *The Quarterly Journal of Economics 133*(2), 553–609.

Poole, K. T. and H. Rosenthal (1984). The polarization of American politics. *The Journal of Politics 46*(4), 1061–1079.

Rajan, R., P. Ramella, and L. Zingales (2023). What purpose do corporations purport? Evidence from letters to shareholders. *NBER Working paper 31054*.

Rouen, E., K. Sachdeva, and A. Yoon (2023). The evolution of ESG reports and the role of voluntary standards. *Working Paper*.

Sautner, Z., L. van Lent, G. Vilkov, and R. Zhang (2023a). Firm-level climate change exposure. *The Journal of Finance 78*(3), 1449–1498.

Sautner, Z., L. van Lent, G. Vilkov, and R. Zhang (2023b). Pricing climate change exposure. *Management Science, forthcoming*.

Zhang, S. (2023). Carbon returns across the globe. *Working paper*.

**Figure 1.** Prosocial Norms Throughout U.S. History

This figure illustrates key historical anecdotes of public discourse about environmental and social issues shaping prosocial norms.

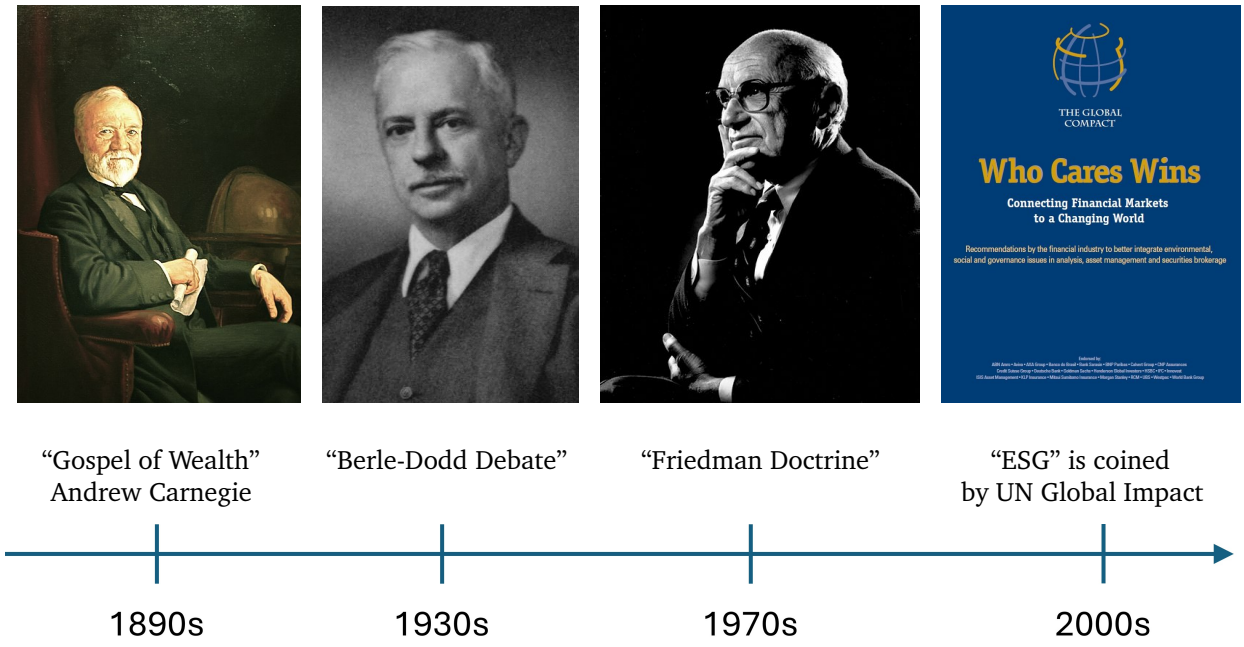
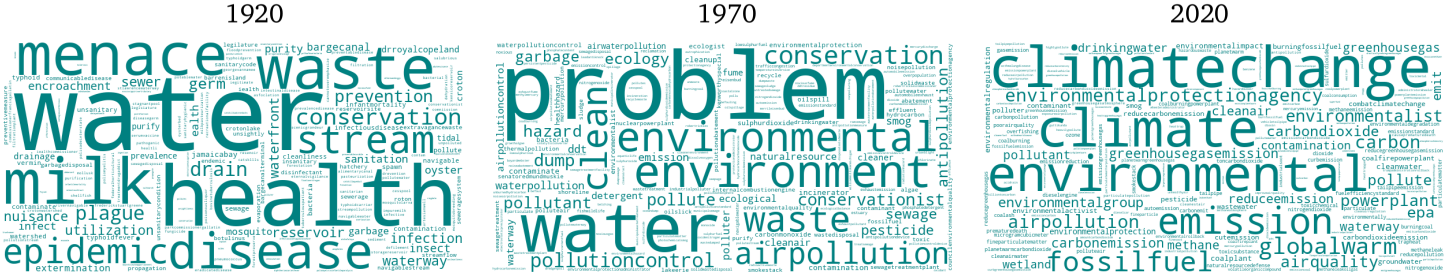


**Figure 2.** Evolving Context of Prosocial Norms

This figure presents word clouds for select years depicting the top 500 keywords in terms of cosine similarity scores weighted by frequency, generated from the seed words "*Pollution*" (Panel A), or "*inequality*" and "*discrimination*" (Panel B), where cosine similarity scores are computed from Word2Vec models trained over the decade prior to each respective year.

Panel A. Keywords related to "*Pollution*"




Panel B. Keywords related to "*Inequality*" and "*Discrimination*"

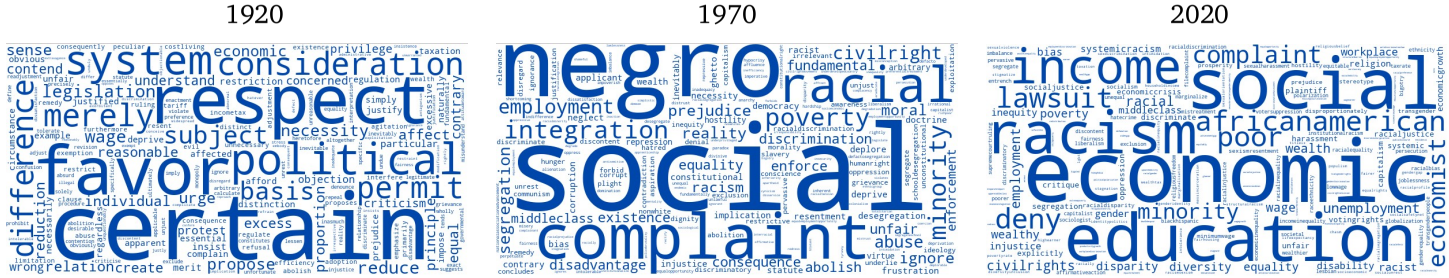


**Figure 3.** Methodology

This figure illustrates our methodology of constructing the Prosocial Norms Index (PNI) by training decade-specific language models on historical news text from the Wall Street Journal (WSJ) and New York Times (NYT), and using the models to span time-varying E&S dictionaries from time-invariant seed words.

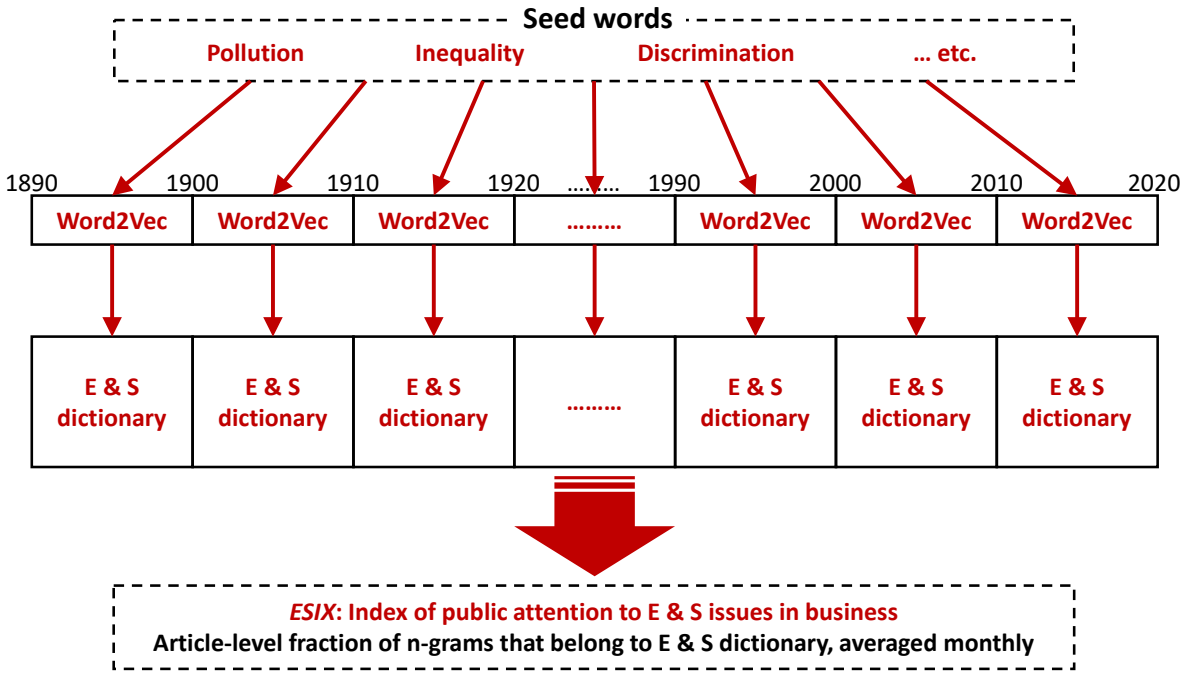
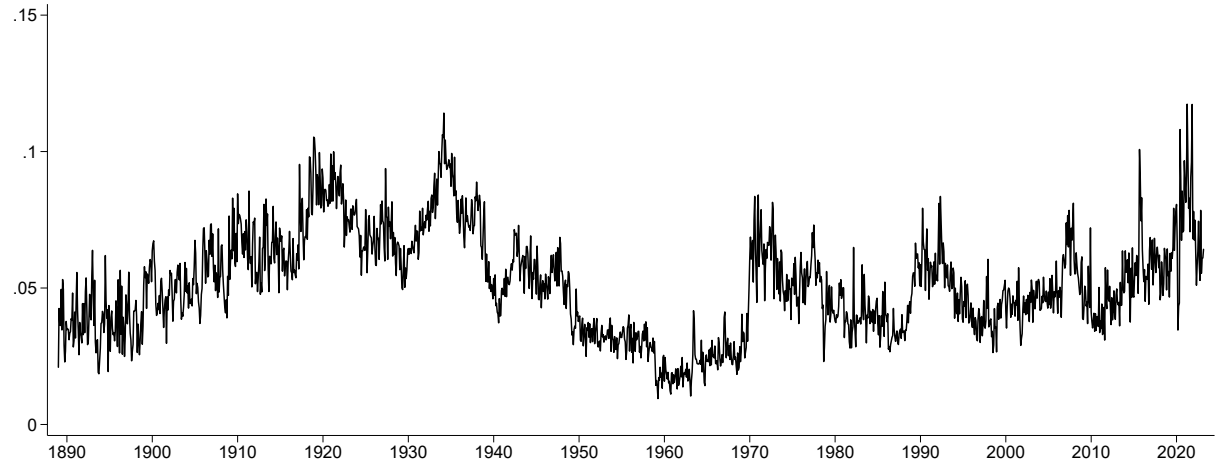


**Figure 4.** The Prosocial Norms Index (PNI)

This figure plots the historical monthly time-series of PNI (Panel A) and its components, PNI-E and PNI-S (Panel B). In Panel B, notable historical events corresponding to spikes or trends in the PNI-E or PNI-S series are annotated in green or red, respectively.

Panel A. Overall PNI




Panel B. PNI-E and PNI-S

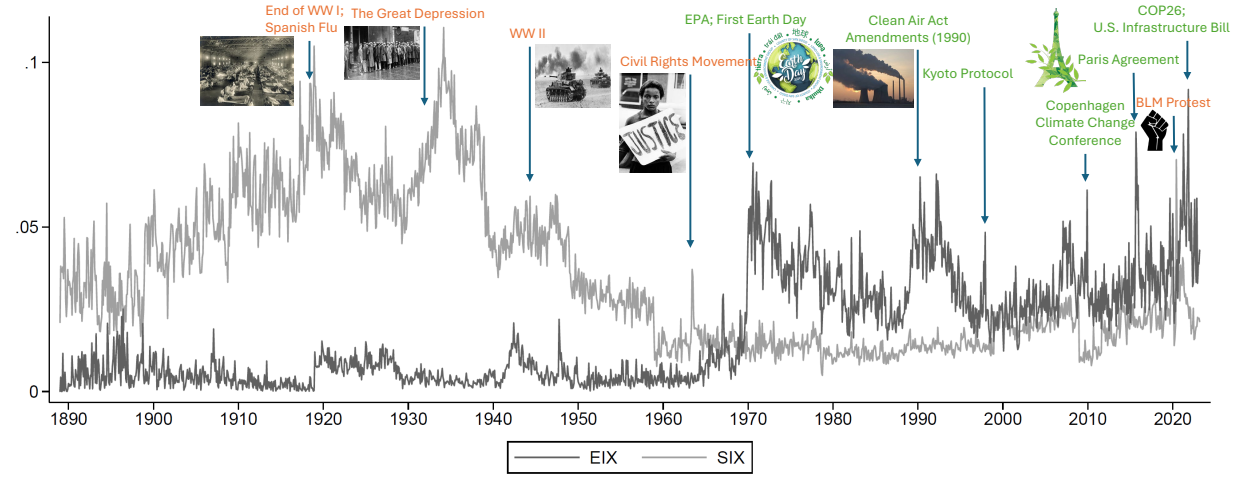
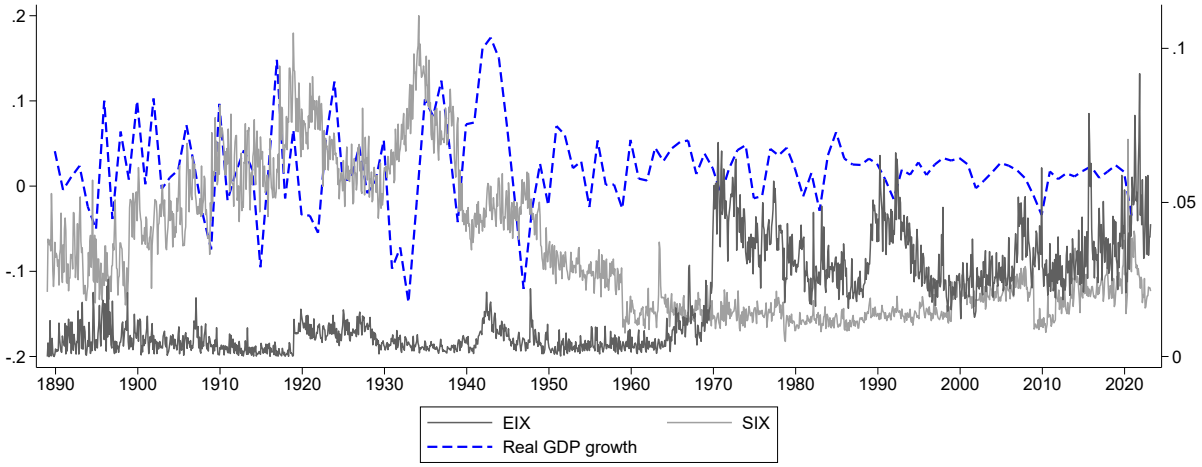


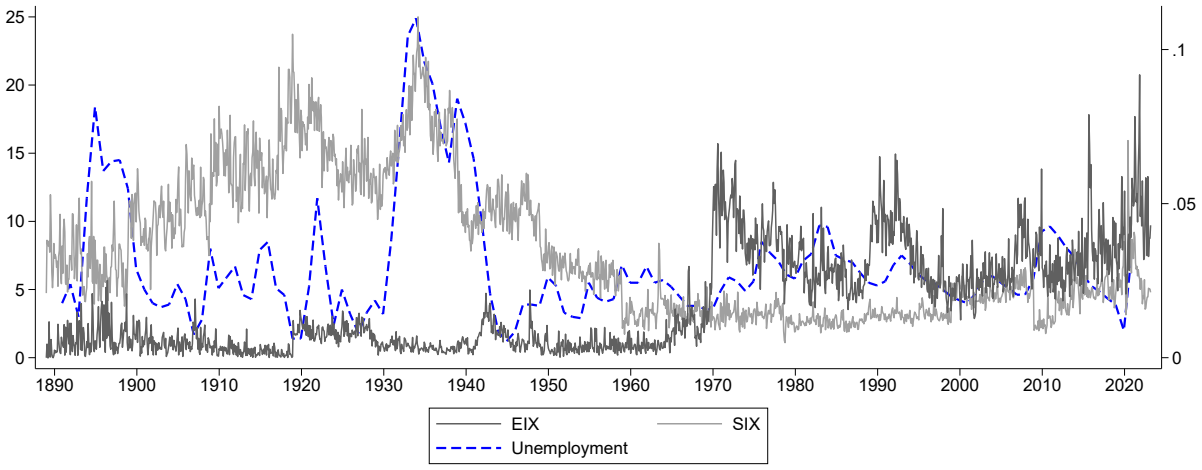
**Figure 5.** PNI and Macroeconomic Conditions

This figure plots the historical monthly time-series of PNI-E and PNI-S, together with annual real GDP growth (Panel A), unemployment (Panel B), NBER recession dummies (Panel C), and wealth inequality (Panel D).

Panel A. Real GDP growth



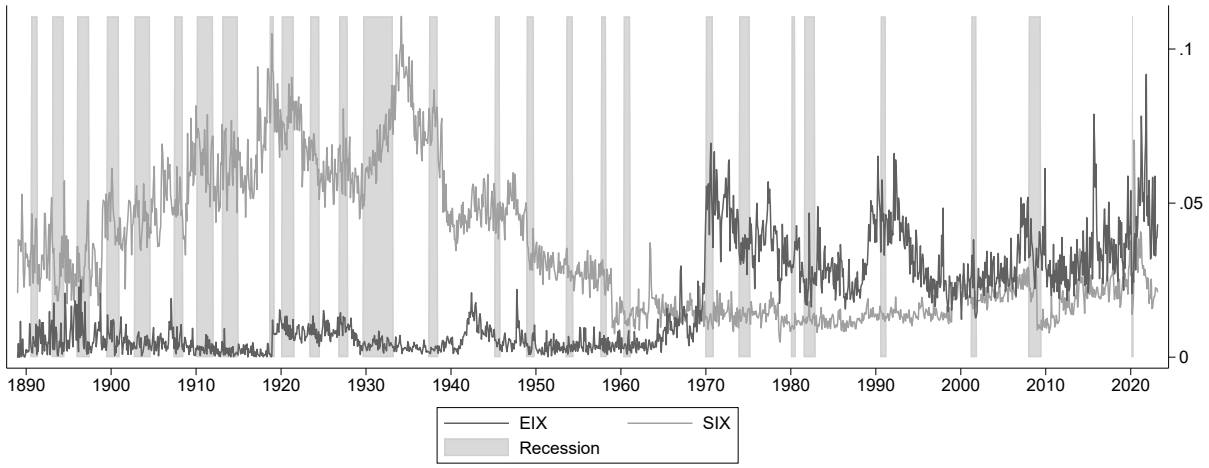

Panel B. Unemployment




*(continued)*

**Figure 5.** PNI and Macroeconomic Conditions (continued)

Panel C. Recessions




Panel D. Wealth inequality

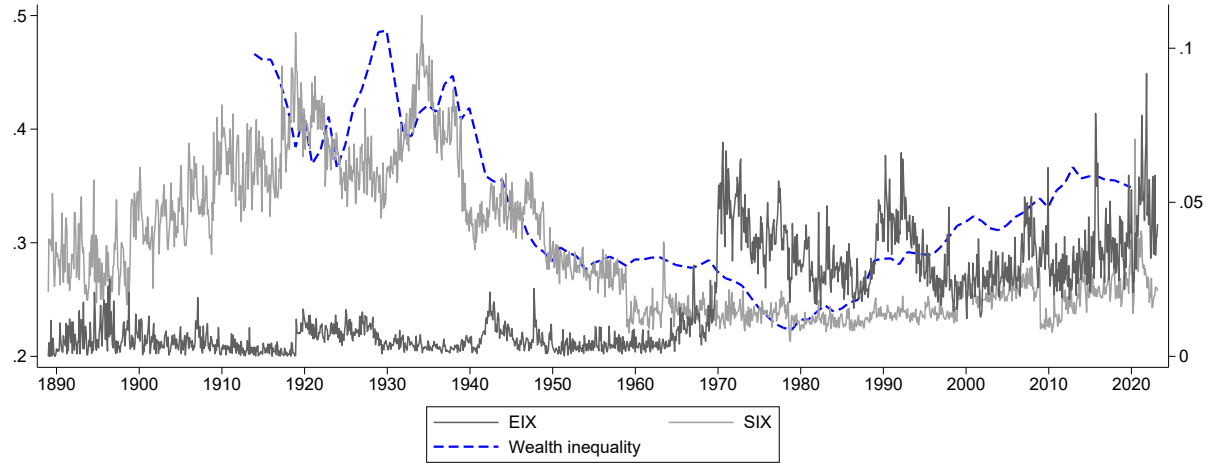
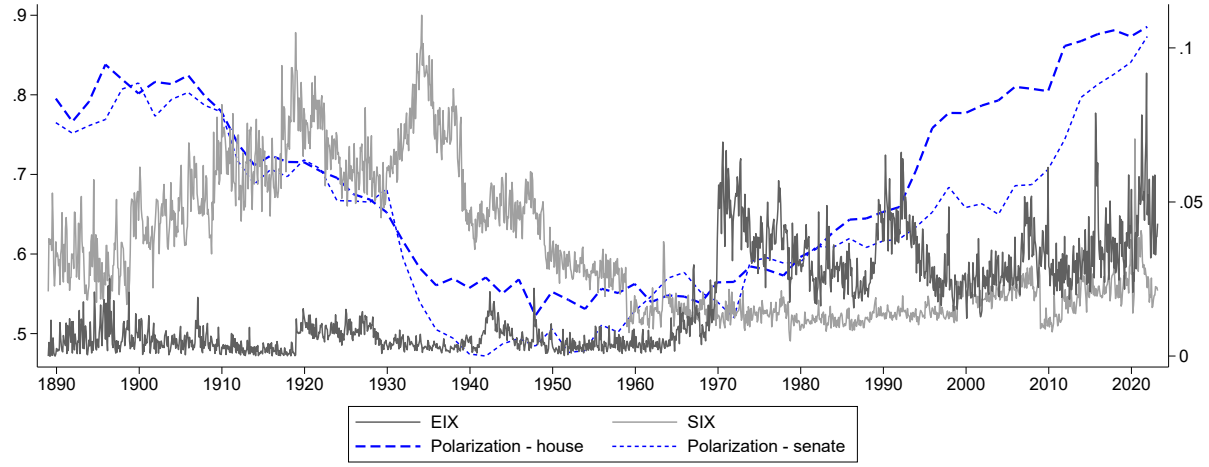


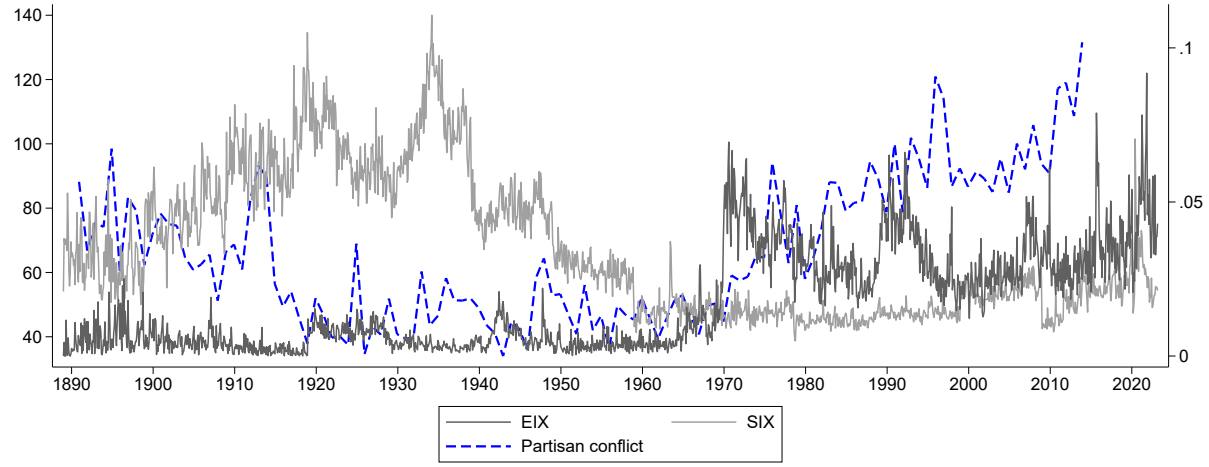
**Figure 6.** PNI and Political Frictions

This figure plots the historical monthly time-series of PNI-E and PNI-S, together with congressional political polarization (Panel A), partisan conflict (Panel B), climate policy uncertainty (Panel C), and geopolitical risk (Panel D).

Panel A. Political polarization



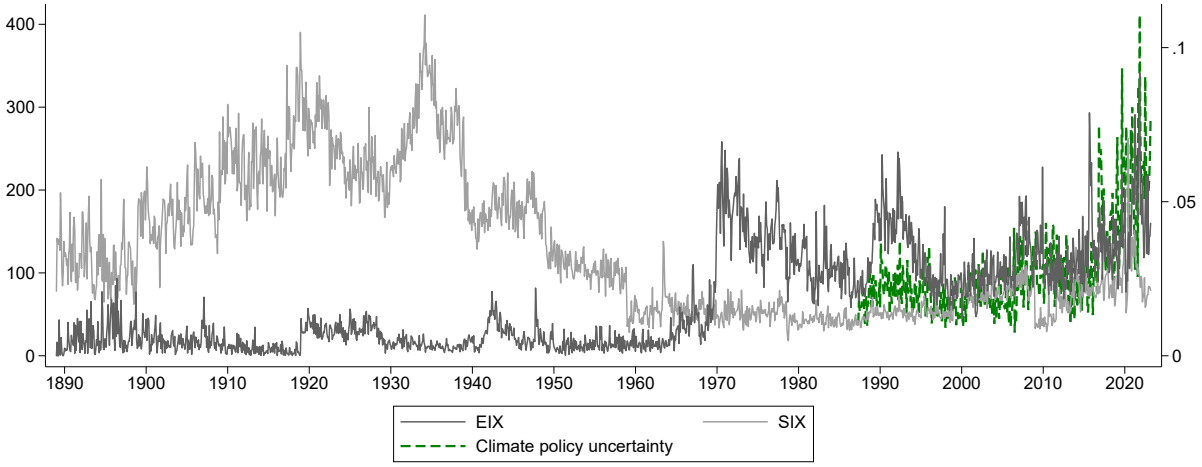

Panel B. Partisan conflict




*(continued)*

**Figure 6.** PNI and Political Frictions (continued)

Panel C. Climate policy uncertainty




Panel D. Geopolitical risk

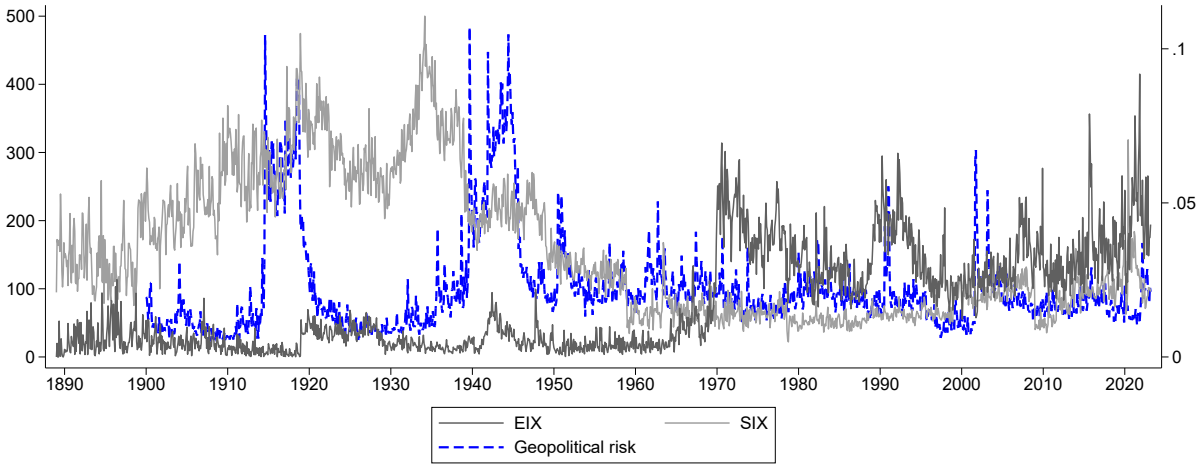
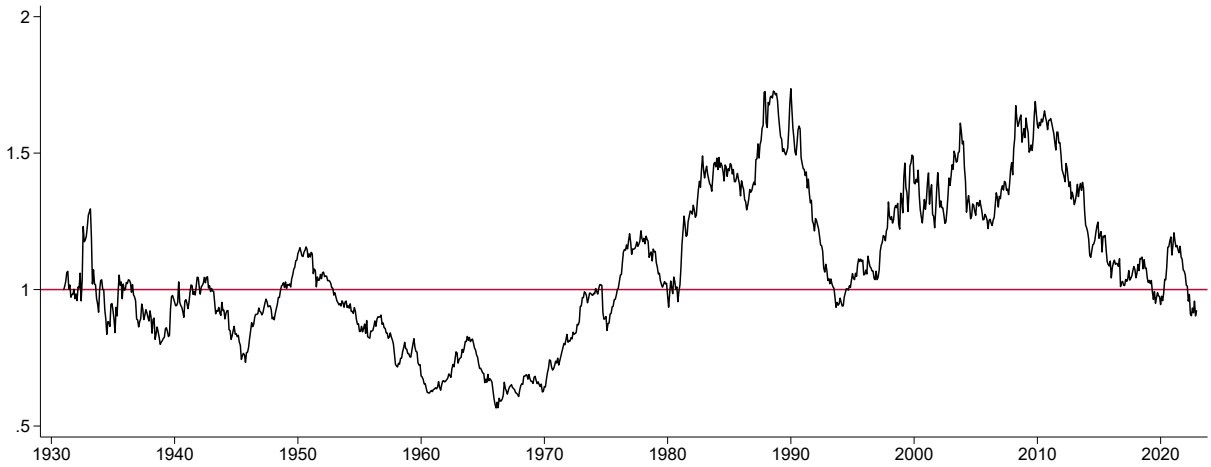


**Figure 7.** The Value of Aligned Values: Exposure to PNI

This figure plots the cumulative return on a portfolio strategy that, each month, longs the top decile of stocks and shorts the bottom decile of stocks sorted on their previous month's PNI exposures. Stock-month level PNI exposure is the PNI-beta computed from five-year rolling window regressions of the stock's return in excess of the risk-free rate, on the Fama and French (1993) three-factor model augmented with monthly innovations in PNI as an additional factor.

Panel A. Overall PNI




Panel B. PNI-E and PNI-S

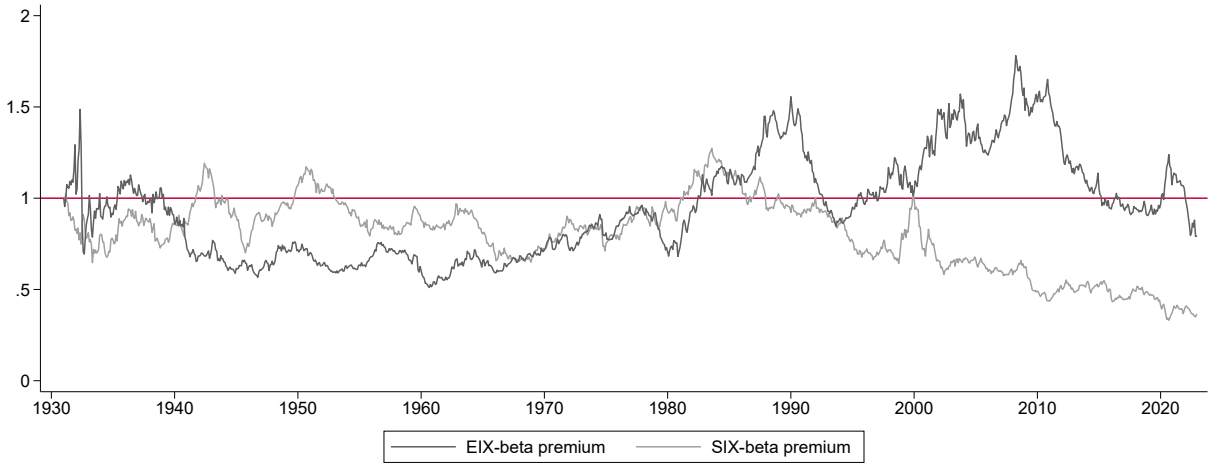
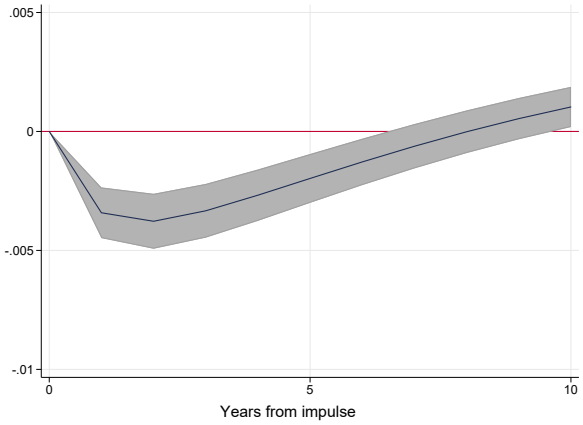


**Figure 8.** Impulse Responses: Corporate Investments

This figure plots orthogonalized impulse response functions from panel vector autoregressions (VARs) using corporate investments (i.e., capital expenditures scaled by lagged assets), yearly average PNI, Tobin's *q*, return on assets (ROA), long-term debt over assets (Debt), and log of total assets (Size) as dependent variables. We fit a multivariate panel regression of each dependent variable on the first lag of itself and on the first lags of all other dependent variables using generalized method of moments (GMM). We control for firm fixed effects by taking first-differences. Standard errors are adjusted for clustering at the firm level. In Panel A, the impulse responses and their 95% confidence intervals are drawn with respect to PNI as the impulse variable and investment as the response variable. In Panel B, the impulse responses and their 95% confidence intervals are drawn with respect to PNI-E or PNI-S as the impulse variable and investment as the response variable.

Panel A. Overall PNI




Panel B. PNI-E and PNI-S

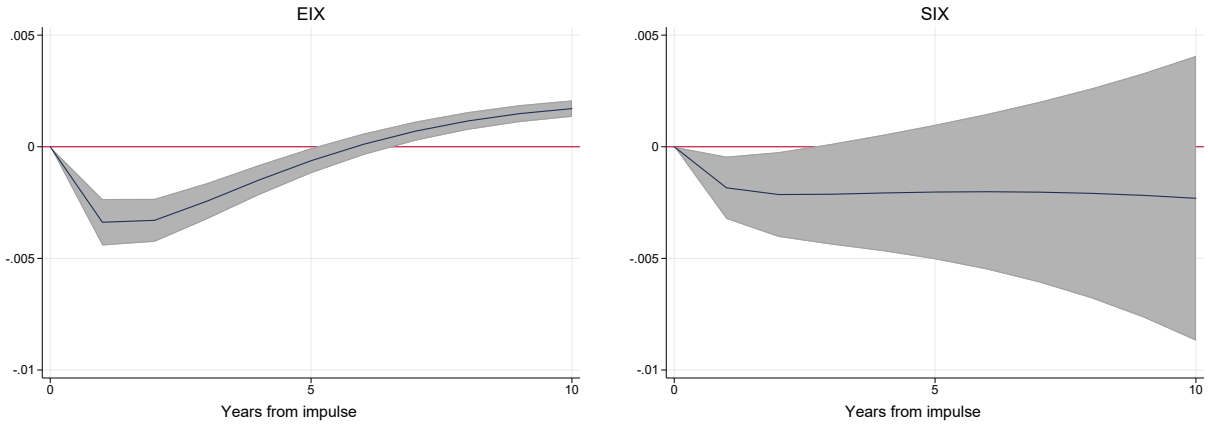


**Figure 9.** Impulse Responses: Corporate Investment Efficiency

This figure plots orthogonalized impulse response functions from panel vector autoregressions (VARs) using corporate investments (i.e., capital expenditures scaled by lagged assets), yearly average PNI, Tobin's $q$, return on assets (ROA), long-term debt over assets (Debt), and log of total assets (Size) as dependent variables. We fit a multivariate panel regression of each dependent variable on the first lag of itself and on the first lags of all other dependent variables using generalized method of moments (GMM). We control for firm fixed effects by taking first-differences. Standard errors are adjusted for clustering at the firm level. The impulse responses and their 95% confidence intervals are drawn with respect to PNI as the impulse variable and investment as the response variable. The left and right charts each plot impulse responses for the subsample of firms with above- or below-median values of Tobin's $q$, respectively, where each firm's Tobin's $q$ is estimated as its earliest available value during its sample period.

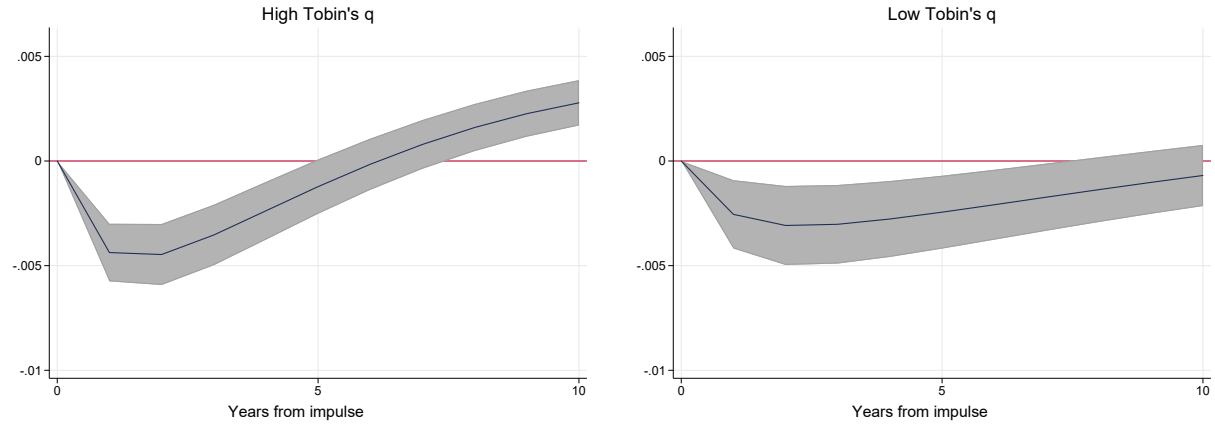
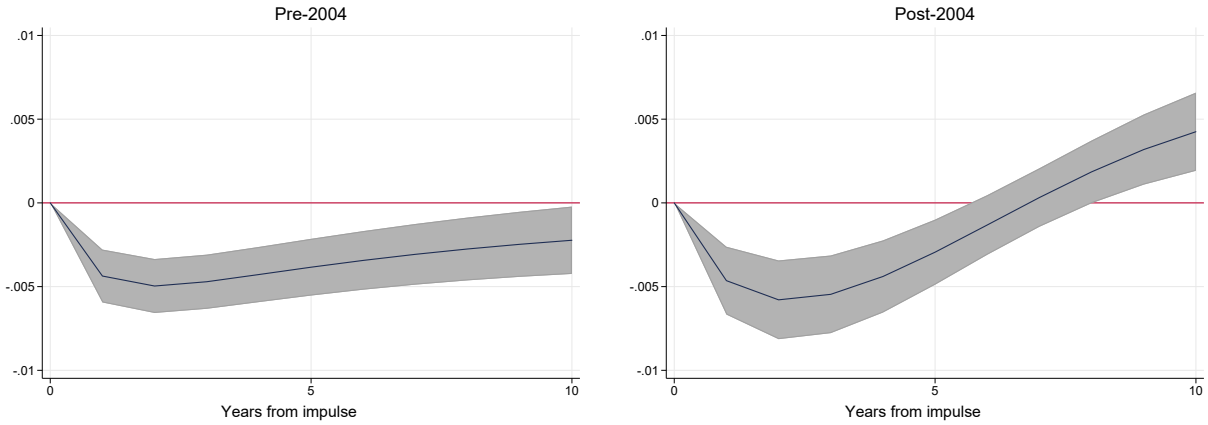


**Figure 10.** Impulse Responses of Corporate Investments: Subsamples

This figure plots orthogonalized impulse response functions from panel vector autoregressions (VARs) using corporate investments (i.e., capital expenditures scaled by lagged assets), yearly average PNI, Tobin's $q$, return on assets (ROA), long-term debt over assets (Debt), and log of total assets (Size) as dependent variables. We fit a multivariate panel regression of each dependent variable on the first lag of itself and on the first lags of all other dependent variables using generalized method of moments (GMM). We control for firm fixed effects by taking first-differences. Standard errors are adjusted for clustering at the firm level. The impulse responses and their 95% confidence intervals are drawn with respect to PNI as the impulse variable and investment as the response variable. Panel A estimates impulse responses for subsamples corresponding to subperiods before and after 2004. Panel B estimates impulse responses for subsamples consisting of firms in the manufacturing / oil & gas / utilities sectors and firms in all other industries based on Fama-French 10 industry classifications.

Panel A. Before and After 2004 (i.e., Emergence of "ESG")




Panel B. Manufacturing / Oil & Gas / Utilities and Others

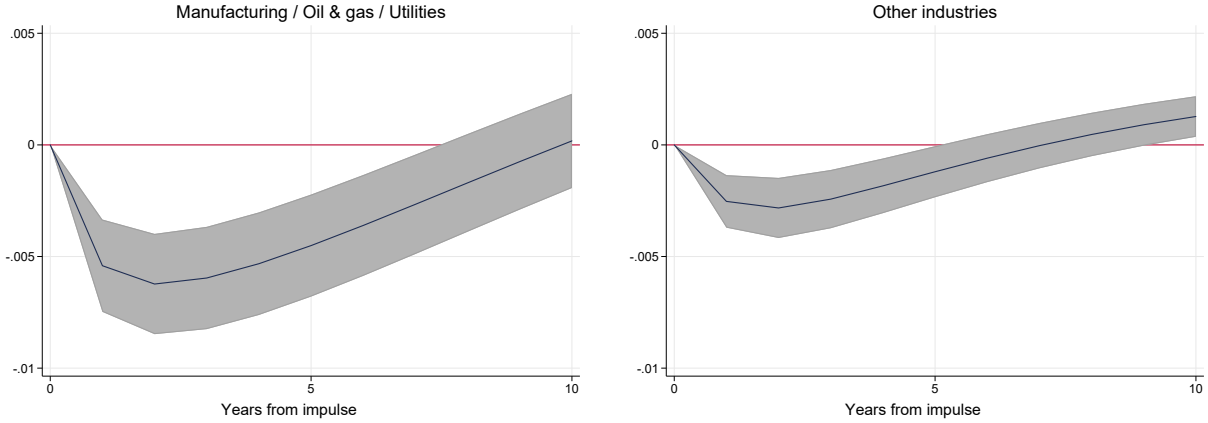


**Table 1.** Summary Statistics

This table presents time-series summary statistics (i.e., Mean, standard deviation, minimum, 25th percentile, median, 75th percentile, maximum) of PNI, PNI-E, and PNI-S. The table also presents minimum and maximum values of PNI, PNI-E, and PNI-S for the pre-1970 and post-1970 subperiods.

| | | | | | | | | Pre-1970 | | Post-1970 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | Std.dev. | Min | 25th | Median | 75th | Max | Min | Max | Min | Max |
| PNI | 0.05 | 0.02 | 0.01 | 0.04 | 0.05 | 0.06 | 0.12 | 0.01 | 0.11 | 0.02 | 0.12 |
| PNI-E | 0.02 | 0.02 | 0.00 | 0.00 | 0.01 | 0.03 | 0.09 | 0.00 | 0.04 | 0.01 | 0.09 |
| PNI-S | 0.03 | 0.02 | 0.00 | 0.02 | 0.03 | 0.05 | 0.11 | 0.01 | 0.11 | 0.00 | 0.07 |

## **Table 2.** PNI and Macroeconomic Conditions

This table presents results from time-series regressions of PNI (Panel A), PNI-E (Panel B), and PNI-S (Panel C) on macroeconomic and political variables. These variables include real GDP growth, unemployment, NBER recession dummies, wealth inequality, congressional polarization at the House and Senate, partisan conflict, climate policy uncertainty, and geopolitical risk. Heteroskedasticity-robust standard errors are reported in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

| | Real GDP growth | Unemployment | Recession | Wealth inequality | Polarization | | Partisan conflict | Climate policy uncertainty | Geopolitical risk |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | House | Senate | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| *Panel A. Dependent variable: PNI* | | | | | | | | | |
| Macro variable | -0.020** | 0.001*** | 0.006*** | 0.172*** | 0.029*** | 0.033*** | -0.000*** | 0.000*** | -0.000 |
| | (0.010) | (0.000) | (0.001) | (0.007) | (0.006) | (0.006) | (0.000) | (0.000) | (0.000) |
| Adj $R^2$ | 0.002 | 0.054 | 0.017 | 0.360 | 0.029 | 0.036 | 0.024 | 0.267 | -0.001 |
| *Panel B. Dependent variable: PNI-E* | | | | | | | | | |
| Macro variable | -0.016** | -0.000*** | -0.007*** | -0.109*** | 0.035*** | 0.027*** | 0.000*** | 0.000*** | -0.000*** |
| | (0.006) | (0.000) | (0.001) | (0.004) | (0.005) | (0.005) | (0.000) | (0.000) | (0.000) |
| Adj $R^2$ | 0.002 | 0.016 | 0.035 | 0.211 | 0.059 | 0.032 | 0.282 | 0.193 | 0.027 |
| *Panel C. Dependent variable: PNI-S* | | | | | | | | | |
| Macro variable | -0.004 | 0.001*** | 0.013*** | 0.281*** | -0.006 | 0.006 | -0.000*** | 0.000*** | 0.000*** |
| | (0.013) | (0.000) | (0.001) | (0.006) | (0.006) | (0.006) | (0.000) | (0.000) | (0.000) |
| Adj $R^2$ | -0.001 | 0.075 | 0.058 | 0.623 | 0.000 | 0.000 | 0.219 | 0.122 | 0.011 |
| Observations | 1,584 | 1,572 | 1,611 | 1,284 | 804 | 804 | 1,488 | 432 | 1,479 |

## **Table 3.** Which Industries are Most Exposed to PNI?

This table reports the three Fama-French 30 industries with the highest and lowest average PNI exposures over five two-decade sub-periods throughout our sample (i.e., 1930–1949, 1950–1969, ... , 2010–present).

| | 1930-1949 | 1950-1969 | 1970-1989 | 1990-2009 | 2010-present |
|---|---|---|---|---|---|
| *Panel A. PNI exposure* | | | | | |
| Positive 1 | Restaurants, hotels, motels | Coal | Restaurants, hotels, motels | Metals & non-metallic mining | Beer & liquor |
| Positive 2 | Transportation | Banking & finance | Automobiles & trucks | Automobiles & trucks | Communication |
| Positive 3 | Personal & business services | Textiles | Wholesale | Personal & business services | Business equipment |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Negative 3 | Textiles | Food | Tobacco | Tobacco | Textiles |
| Negative 2 | Business supplies & containers | Aircraft, ships, & railroad | Coal | Healthcare | Tobacco |
| Negative 1 | Healthcare | Printing & publishing | Beer & liquor | Coal | Coal |
| *Panel B. PNI-E exposure* | | | | | |
| Positive 1 | Coal | Restaurants, hotels, motels | Wholesale | Restaurants, hotels, motels | Beer & liquor |
| Positive 2 | Recreation | Communication | Restaurants, hotels, motels | Metals & non-metallic mining | Communication |
| Positive 3 | Utilities | Metals & non-metallic mining | Automobiles & trucks | Recreation | Business equipment |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Negative 3 | Retail | Printing & publishing | Beer & liquor | Electrical equipment | Steel works |
| Negative 2 | Business supplies & containers | Coal | Coal | Healthcare | Tobacco |
| Negative 1 | Beer & liquor | Business equipment | Tobacco | Coal | Coal |
| *Panel C. PNI-S exposure* | | | | | |
| Positive 1 | Restaurants, hotels, motels | Coal | Petroleum & natural gas | Beer & liquor | Tobacco |
| Positive 2 | Transportation | Banking & finance | Restaurants, hotels, motels | Textiles | Healthcare |
| Positive 3 | Personal & business services | Electrical equipment | Metals & non-metallic mining | Communication | Printing & publishing |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Negative 3 | Business supplies & containers | Steel works | Beer & liquor | Construction | Apparel |
| Negative 2 | Textiles | Food | Banking & finance | Tobacco | Food |
| Negative 1 | Healthcare | Aircraft, ships, & railroad | Communication | Coal | Textiles |

## **Table 4.** PNI and Stock Returns

This table presents results from regressions of monthly stock returns on lagged values of PNI/PNI-E/PNI-S exposures. Exposures are obtained from five-year rolling regressions of the stock's return in excess of the risk-free rate, on the Fama and French (1993) three-factor model augmented with monthly innovations in PNI as an additional factor. Panel A uses the raw exposures as continuous explanatory variables. Panel B uses dummy variables indicating whether exposure is positive as explanatory variables. Regressions are estimated on the full sample period from February 1931 to January 2023, and also on separate subsamples for the pre-1970 and post-1970 periods. Standard control variables are included, including lagged values of firm size (i.e., market capitalization), lagged monthly stock return, past twelve months' returns skipping a month, past three year's monthly returns skipping a year, idiosyncratic volatility computed from daily residuals of the Fama and French (1993) three-factor model estimated over the previous twelve months, and the average bid-ask spread. Stock and month fixed effects are included. Standard errors are clustered at the stock level (*** p<0.01, ** p<0.05, * p<0.1).

*Panel A. Continuous exposures*

| | Dependent variable: Monthly stock return | | | | | |
|---|---|---|---|---|---|---|
| | Full sample period | | Pre-1970 | | Post-1970 | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *PNI* exposure | -0.002 | | -0.001 | | -0.002 | |
| | (0.002) | | (0.003) | | (0.003) | |
| *PNI-E* exposure | | -0.000 | | -0.063*** | | 0.001 |
| | | (0.003) | | (0.018) | | (0.003) |
| *PNI-S* exposure | | -0.007*** | | 0.001 | | -0.009*** |
| | | (0.002) | | (0.003) | | (0.002) |
| Observations | 1,817,883 | 1,817,883 | 246,322 | 246,322 | 1,571,559 | 1,571,559 |
| Stock FE | Y | Y | Y | Y | Y | Y |
| Year-by-month FE | Y | Y | Y | Y | Y | Y |
| Stock controls | Y | Y | Y | Y | Y | Y |
| Adj $R^2$ | 0.136 | 0.136 | 0.390 | 0.390 | 0.109 | 0.109 |

*Panel B. Positive exposure dummies*

| | Dependent variable: Monthly stock return | | | | | |
|---|---|---|---|---|---|---|
| | Full sample period | | Pre-1970 | | Post-1970 | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **I**(*PNI* exposure>0) | -0.000 | | -0.000 | | -0.000 | |
| | (0.000) | | (0.001) | | (0.000) | |
| **I**(*PNI-E* exposure>0) | | -0.000 | | -0.002*** | | -0.000 |
| | | (0.000) | | (0.001) | | (0.000) |
| **I**(*PNI-S* exposure>0) | | -0.001** | | 0.000 | | -0.001** |
| | | (0.000) | | (0.001) | | (0.000) |
| Observations | 1,818,367 | 1,818,367 | 246,806 | 246,806 | 1,571,559 | 1,571,559 |
| Stock FE | Y | Y | Y | Y | Y | Y |
| Year-by-month FE | Y | Y | Y | Y | Y | Y |
| Stock controls | Y | Y | Y | Y | Y | Y |
| Adj $R^2$ | 0.136 | 0.136 | 0.390 | 0.390 | 0.109 | 0.109 |

## **Table 5.** PNI and Corporate Investments

This table presents results from firm-year panel regressions of corporate investments (i.e., capital expenditures scaled by lagged assets). In Panel A, investments are regressed on lagged values of yearly average PNI, PNI-E, PNI-S, and the yearly changes (denoted by $\Delta$) in PNI, PNI-E, and PNI-S. Control variables include lagged values of Tobin's *q*, return on assets (ROA), long-term debt over assets (Debt), log of total assets (Size), as well as firm and industry-by-decade fixed effects. In Panel B, PNI, PNI-E, PNI-S, $\Delta$PNI, $\Delta$PNI-E, and $\Delta$PNI-S are interacted with lagged Tobin's *q*, and industry-by-decade fixed effects are replaced by industry-by-year fixed effects. Standard errors are clustered at the firm level (*** p<0.01, ** p<0.05, * p<0.1).

*Panel A. PNI and corporate investments*

| | Dependent variable: Capital expenditures$_t$/Assets$_{t-1}$ | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *PNI* | -0.085*** | | | | -0.081*** | |
| | (0.025) | | | | (0.025) | |
| $\Delta$*PNI* | | -0.027* | | | -0.018 | |
| | | (0.017) | | | (0.017) | |
| *PNI-E* | | | -0.098*** | | | -0.096*** |
| | | | (0.029) | | | (0.030) |
| *PNI-S* | | | -0.015 | | | -0.004 |
| | | | (0.084) | | | (0.090) |
| $\Delta$*PNI-E* | | | | -0.028* | | -0.015 |
| | | | | (0.017) | | (0.017) |
| $\Delta$*PNI-S* | | | | -0.026 | | -0.033 |
| | | | | (0.033) | | (0.036) |
| Tobin's *q* | 0.014*** | 0.014*** | 0.014*** | 0.014*** | 0.014*** | 0.014*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| ROA | 0.024*** | 0.024*** | 0.024*** | 0.024*** | 0.024*** | 0.024*** |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| Debt | -0.031*** | -0.032*** | -0.031*** | -0.032*** | -0.031*** | -0.031*** |
| | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) |
| Size | -0.017*** | -0.017*** | -0.017*** | -0.017*** | -0.017*** | -0.017*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Observations | 202,664 | 202,664 | 202,664 | 202,664 | 202,664 | 202,664 |
| Firm FE | Y | Y | Y | Y | Y | Y |
| Industry-by-decade FE | Y | Y | Y | Y | Y | Y |
| Adj $R^2$ | 0.420 | 0.420 | 0.420 | 0.420 | 0.420 | 0.420 |

*(continued)*

**Table 5.** PNI and Corporate Investments (continued)

*Panel B. PNI and investment-q sensitivity*

| | Dependent variable: Capital expenditures$_t$/Assets$_{t-1}$ | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *PNI* × Tobin's *q* | -0.085*** | | | | -0.070*** | |
| | (0.018) | | | | (0.018) | |
| Δ*PNI* × Tobin's *q* | | -0.073*** | | | -0.058** | |
| | | (0.027) | | | (0.027) | |
| *PNI-E* × Tobin's *q* | | | 0.072** | | | 0.094*** |
| | | | (0.030) | | | (0.032) |
| *PNI-S* × Tobin's *q* | | | -0.517*** | | | -0.543*** |
| | | | (0.057) | | | (0.077) |
| Δ*PNI-E* × Tobin's *q* | | | | -0.066** | | -0.057** |
| | | | | (0.026) | | (0.026) |
| Δ*PNI-S* × Tobin's *q* | | | | -0.096** | | 0.056 |
| | | | | (0.040) | | (0.055) |
| Tobin's *q* | 0.018*** | 0.014*** | 0.021*** | 0.014*** | 0.017*** | 0.021*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| ROA | 0.021*** | 0.021*** | 0.021*** | 0.021*** | 0.021*** | 0.021*** |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Debt | -0.028*** | -0.028*** | -0.027*** | -0.028*** | -0.028*** | -0.027*** |
| | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) | (0.006) |
| Size | -0.018*** | -0.018*** | -0.018*** | -0.018*** | -0.018*** | -0.018*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Observations | 202,657 | 202,657 | 202,657 | 202,657 | 202,657 | 202,657 |
| Firm FE | Y | Y | Y | Y | Y | Y |
| Industry-by-year FE | Y | Y | Y | Y | Y | Y |
| Adj $R^2$ | 0.437 | 0.437 | 0.437 | 0.437 | 0.437 | 0.437 |