Measuring the Local and Global Impact of Corporate Social Responsibility Using Artificial Intelligence

Dina El Mahdy[*], Denis Gracanin[**], Rachel Soloveichik[***], Majid Behravan[*], Sezer Dumen[**]

Morgan State University[*], Virginia Tech[**], Bureau of Economic Analysis[***]

ABSTRACT

This study introduces an AI-powered methodology that uses Natural Language Processing (NLP) and Large Language Models (LLMs) to construct a comprehensive database of U.S. firms' corporate social responsibility (CSR) activities from 2000–2025. The database extracts and standardizes nearly 300,000 observations, including 91,395 firm-donation-year records, from 1.2 billion news articles, SEC filings, sustainability reports, and social media. The approach also enables the valuation of non-monetary contributions, including approximately $6 billion in employee volunteering across 14,990 records, using replacement-cost methods. Preliminary analysis reveals that CSR donations increased from under $1 billion in 2000 to over $2.5 billion by 2025, driven in part by a post-2018 surge in non-monetary activities. Contributions are highly concentrated, with firms in the pharmaceutical, telecommunications, utilities, and financial sectors accounting for more than 70% of total donations. Statistical analysis further shows that donations decline significantly with distance (i.e., geographic proximity to the recipient), especially for foundation-backed firms and for local categories such as health and education. By overcoming longstanding fragmentation in CSR data, this scalable approach enables more robust empirical analysis of key drivers, including firm size, governance, and market power. A forthcoming public database is expected to inform GDP adjustments, national accounting, and policy design.

JEL Classification: C55, C81, D22, M14, O47

Keywords: economic measurement, corporate social responsibility (CSR), charitable donations, non-monetary contributions, in-house nonprofits, environmental, social, and governance (ESG), machine learning, artificial intelligence, national accounting, GDP, stakeholder theory, geographic proximity

[*]**Corresponding Author:** Department of Accounting and Finance, Earl. G. Graves School of Business and Management. 4100 E. Cold Spring Lane, 21218. Email: dina.elmahdy@morgan.edu.

Acknowledgement: This material is based upon work supported by the National Science Foundation under Award No. [2401824]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

I. INTRODUCTION

Corporate Social Responsibility (CSR) has been defined as the firm's responsibility for its impact on society (European Commission 2011). It is closely related to environmental, social, and governance (ESG) activities that extend beyond shareholder profit maximization. Despite its societal importance, existing measures of CSR are often fragmented and inconsistent. This is because relevant information is dispersed across multiple sources—such as regulatory filings, news articles, social media, and firm announcements—and there is no widely accepted standard for reporting. As a result, most prior research that has addressed CSR uses databases that have been criticized for their weak internal and construct validities (Raghunandan and Rajgopal 2020).

The accounting community spends more than $1 trillion annually summarizing complex financial data into easy-to-use measures (Soloveichik 2023; Ewans et al. 2021). If similar easy-to-use measures were available for corporate social responsibility (CSR), then those stakeholders could make decisions based on the CSR aspects of a project.

Following research trends (Splechtna et al. 2023), this paper develops a consistent and organized CSR database using artificial intelligence (AI). More specifically, this study constructs CSR database using AI-driven extraction frameworks that combine web-scale news processing with multi-source disclosure analysis. The web-scale news processing began with a universe of 1.2 billion news articles published between 2000 and 2025 that were screened to yield 1.1 million unique articles with accessible stored texts. The multi-source disclosure analysis began with a broad universe of regulatory filings, corporate sustainability reports, and social media disclosures. To extract CSR activity from these sources, a keyword-anchored context window strategy was applied, and the resulting text window was analyzed using an LLM instructed to record CSR activity only when a voluntary transfer of value is present and the recipient qualifies as a social or

community organization. Each extracted record is subsequently passed through a second validation stage to remove false positives. This dual-stage extraction and validation process substantially reduces measurement error and ensures that retained observations reflect genuine CSR activity. In total, the database used in this paper contains approximately 300,000 individual CSR observations spanning 2000-2025 and includes 91,395 observations for firms that can be linked with Compustat and other sources.

In-house CSR activity is difficult to value because activity is typically reported in narrative rather than monetary terms. Therefore, this study develops a systematic approach grounded in established nonprofit valuation frameworks (Salamon et al. 2011) for fee waivers, employee volunteering hours, in-kind product donations, facility access, and training programs. Preliminary analysis reveals that total in-house CSR activity had a value of $6 billion in the 14,990 records analyzed so far, that the finance sector dominated with a 25 percent share of CSR activity, and that employee volunteering accelerated after 2018. Measured gross domestic product (GDP) would increase slightly if this in-house CSR activity were tracked consistently with cash donations to nonprofits.

In addition to building a comprehensive CSR database, this study also examines firms' activities with a particular focus on the geographic allocation of charitable donations.

Drawing on the broader literature that documents the role of proximity in information access, monitoring, strategic behavior, and decision-making (e.g., Choi et al. 2012; Jennings et al. 2017; Lai et al. 2020; Deng et al. 2020; Huang et al. 2022), this study investigates whether firms' donation decisions are systematically related to the geographic proximity of the causes they support. The study addresses several interrelated research questions, including whether decisions associated with CSR have evolved over time, which types of firms engage in CSR activities that benefit their local community rather than global causes, and why firms initiate CSR activities more broadly. By

addressing these questions, the study aims to enhance understanding of the determinants, evolution, and geographic orientation of corporate charitable giving.

The remainder of this paper is organized as follows. Section two discusses literature review and hypothesis development. Section three provides the research methodology and steps for constructing the CSR database. Section four provides preliminary summary statistics and results. This paper then concludes with plans for future research and plans for database posting in section five. Finally, Appendix A provides an example of the CSR activities studied, Appendix B lists variable definitions, and Appendix C provides figures from the preliminary analysis.

II. LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Why Do Firms Engage in CSR Initiatives?

Neoclassical economic theory—often associated with Milton Friedman—posits that a firm's primary responsibility is to maximize shareholder profit. From this perspective, CSR is viewed as a costly endeavor with uncertain economic returns. For example, CSR initiatives may involve offering reduced prices to underserved customers, maintaining green spaces in economically disadvantaged areas, or making charitable donations. Such activities can impose significant financial burdens, potentially weaken a firm's economic performance (Jensen 2002) and place it at a competitive disadvantage relative to firms that do not engage in CSR (Aupperle, Carroll, and Hatfield 1985).

Building on this perspective, neoclassical theory does recognize that certain forms of CSR can be economically rational. For instance, investments in diversity or adherence to environmental regulations can reduce the risk of litigation and associated costs (McGuire, Sundgren, and Schneeweis 1988). CSR can also enhance productivity if it improves employee morale or generates valuable brand equity when customers perceive the firm's actions as socially beneficial. Moreover,

CSR can be seen as a firm's private response to market imperfections—such as public goods and bads, altruism, imperfect competition, or incomplete contracts—where proactive engagement can create long-term economic advantages.

Furthermore, Modern stakeholder theory holds that the firm's responsibilities extend beyond maximizing shareholders' value (Cornell and Shapiro 1987). Underlying this theory, the firm is responsible for satisfying various stakeholders who have either implicit or explicit claims. An example of an implicit claim is customer expectation of high product quality, and an example of an explicit claim is bondholder demand. It is assumed that implicit claims are more flexible than explicit claims, and hence less risky. An important implication of implicit claims is that stakeholders may demand that they be transformed into explicit claims if the firm does not put sufficient focus on satisfying implicit claims. Underlying this notion, CSR enables the firm to signal focus and therefore keep some claims as flexible implicit claims. In other words, lower investment in CSR may protect the firm from certain types of risk (McGuire et al. 1988).

Decisions to engage in CSR can be influenced by several factors. The stakeholder theory predicts that firms face pressure from a diverse group with conflicting interests. Therefore, firms target activities that strike a balance among stakeholders' needs without causing detriment to current shareholders' demand for profit maximization. For example, some industries may arrange for their compliance with mandatory environmental laws to be combined with high profile CSR activities. Pagano and Volpin (2005) argue that firms engage in CSR as a self-defense mechanism to avoid hostile takeovers. McGuire et al. (1988) find that determinants of CSR include the firm's prior financial performance. Therefore, firms with high past performance are willing to devote slack resources to CSR activities. Another stream of research examined the association between CSR

and fraud. For example, Hemingway and Maclagan (2004) claim that firms engage in CSR just to cover corporate misconduct.

Because economists often view CSR as a 'catch-all' term that involves various concepts (Benabou and Tirole 2010), firms can engage in CSR for various reasons:

a) Long-term perspective philanthropy, also known as doing well by doing good. Firms may also use CSR to meet societal demands (Crifo and Forget 2015) such as the need for economic growth (Margolis et al. 2009). This notion assumes that managers can identify CSR activity that doesn't require sacrificing short-term profitability.

b) Stakeholder-delegated philanthropy. Firms give strategically to the community (Saiia et al., 2003) and increasingly integrate voluntary reporting of social, governance, and environmental factors into business operations and interactions with stakeholders in an attempt to balance planet, people, and profit (Van Marrewijk 2003). Firms engage in CSR under this scenario based on pressure from various stakeholders of the firm (employees, suppliers, investors). In this case, the firm is viewed as the vehicle through which stakeholders express their values and sacrifice some wealth to achieve social good.

c) Insider-initiated corporate philanthropy. In this case, CSR initiatives are driven by insider demand for socially responsible behavior. This demand could come from the management or board of directors, for example. Other stakeholders may also pressure the firm to engage in CSR (Perez-Batres et al. 2012). To the extent that these decisions are highly contagious, board interlocks and other social networks become an important determinant/factor in these decisions. It would be interesting to examine whether firms donate to charities that their board members and other stakeholders are involved in and to what causes they contribute the most.

Understanding why companies engage in community contributions is important for multiple reasons. For investors, it is critical to determine whether socially responsible activities provide genuine financial benefits. Ignoring this could undermine investor wealth, contrary to the principles of firm and agency theory. Equally important is verifying the authenticity of a firm's reported CSR activities. For example, misreporting—commonly referred to as greenwashing—can result in financial loss while providing no real societal benefit. Given the conflicting evidence on firms' responsibilities to stakeholders, analyzing the motives behind CSR and the mechanisms through which it affects society is essential.

The database introduced in the next section addresses these challenges by providing a comprehensive, structured source of CSR data. It will enable researchers to trace the flow of CSR resources across industries, quantify intangible corporate contributions, and rigorously evaluate the societal and financial impact of CSR initiatives. This study uses a preliminary version of the database to test a specific CSR hypothesis, while the final database, which will be made publicly available online, can support a wide range of future research inquiries. Beyond academia, national accountants can incorporate CSR data into measures of economic growth, capturing previously untracked corporate contributions. Policymakers, government officials, managers, and investors will also benefit from access to reliable and detailed CSR information, allowing them to make informed decisions and strategies related to social responsibility.

The Controversy around CSR Measurement

Currently, no uniform standards govern CSR disclosure.[1] Because of this lack of uniform standards, 36% of executives mentioned that they do not track their sustainability efforts in a 2022

[1] For example, the United Nations General Assembly established a set of 17 global sustainable development goals, aimed at addressing various global challenges. Some firms use those goals as a framework to plan and communicate their CSR activities (Fatima and Elbanna 2022). Several voluntary CSR reporting frameworks that firms can use

Harris Poll survey (Peters 2022). The same survey reports that 68% of U.S. firms' executives admitted to misreporting CSR activity. CSR data often involves financial and non-financial metrics, which require subjective assessment that may introduce bias. Some companies may only disclose positive CSR impact as a greenwashing mechanism. Alternatively, some companies may exaggerate their positive CSR impact. CSR data has been operationalized in empirical research using indicator values for the presence or absence of CSR, or rating scores that lack verifiability and comparability. Anecdotal evidence suggests that these ratings tend to be subjective and incomplete (Krappel et al. 2021). Additionally, not all firms are covered by these ratings agencies, and/or different agencies can give conflicting ratings to the same firm.

Our methodology is unique because it will not only provide monetary CSR donations but also quantify non-monetary CSR activity that can influence the overall well-being of society. Our dataset incorporates vital non-financial data points, such as CSR strategy, which signifies a firm's dedication to sustainability. Our study extends the limited research that employs Machine Learning in CSR studies (e.g., Teoh et al., 2019) and includes previously ignored elements of CSR related to human rights and focused on financial variables (Gao and Zhang 2015). Our research also responds to prior literature (Cracolici et al. 2010) that calls for a novel approach to measure and analyze national and regional GDP by expanding the national accounting framework to include non-monetary CSR activities as a type of in-house nonprofit.

Corporate Donations and Geographic Proximity

The literature on geographic proximity (e.g., Huang et al. 2022; Lai et al. 2020; Deng et al. 2020; Jennings et al. 2017; Choi et al. 2012) documents its relevance for decision-making in economic,

include the Global Reporting Initiative (GRI), the Sustainability Accounting Standards Board (SASB), the Carbon Disclosure Project (CDP), and the Task Force on Climate-related Financial Disclosures (TCFD).

accounting, and finance settings. Prior studies show that geographic proximity can shape outcomes through mechanisms such as home bias, reduced information costs, enhanced knowledge sharing, and improved monitoring. These effects have been linked to a wide range of financial phenomena, including earnings management, audit quality, corporate misconduct, and analyst behavior. However, the extent to which these mechanisms generalize to firms' CSR activity remains unclear.

Lai et al. (2020) find that CEOs with strong ties to their place of origin exhibit less short-term focused decision-making. Similarly, Jennings et al. (2017) show that geographic proximity among firms within an industry lowers analysts' information costs and enhances peer knowledge. While such findings suggest that geographic attachment may influence certain strategic corporate decisions, CSR activity could differ fundamentally from operational or investment choices. After all, CSR decisions are often guided by reputational concerns, stakeholder expectations, and firm-wide strategies that may prioritize visibility or thematic alignment over geographic proximity. Firms may therefore face limited incentives to favor geographically proximate causes when making philanthropic allocations.

Evidence from the auditing literature also highlights the nuanced role of proximity. Choi et al. (2012) and Huang et al. (2022) document that local auditors benefit from informational advantages that improve audit quality. However, Deng et al. (2020) caution that proximity can introduce countervailing forces, such as favoritism or political influence, which may weaken monitoring effectiveness. These findings suggest that geographic proximity does not uniformly improve outcomes and may, in some contexts, have neutral or offsetting effects. In the context of charitable giving, similar countervailing considerations—such as centralized CSR governance, global branding strategies, or regulatory scrutiny—may limit the role of local proximity.

Taking together, the existing literature indicates that while geographic proximity can influence decision-making in settings characterized by information asymmetry and monitoring needs, its relevance for corporate charitable donations is ambiguous. Charitable giving decisions may be sufficiently standardized, strategically centralized, or reputationally motivated that firms do not systematically favor geographically proximate causes. Accordingly, despite documented proximity effects in other economic contexts, we posit that firms' charitable donations are not associated with the geographic proximity of the causes they support. Based on this reasoning, we formulate the following null hypothesis:

H0: There is no association between a firm's charitable donations and geographic proximity.

III. RESEARCH METHODOLOGY

Measuring CSR consistently at scale requires AI-based methods. Therefore, AI tools, such as NLP can be used to automate the process of extracting CSR data from various sources. These tools are particularly useful because CSR initiatives are scattered across various documents, websites, and resources, predominantly on social media (Holcomb, Upchurch, and Okumus, 2007). Appendix A provides examples of some CSR activities extracted and organized using NLP. Next, AI tools can enhance the consistency and comparability of CSR data by creating a standardized format for CSR that adjusts for sources of bias such as sentiment and tone. These adjustments are made by identifying patterns across key data points and converting unstructured reporting into a standardized and less biased format. Finally, AI tools enable data analysis through either visualization (e.g., maps) and advanced statistical methods.

Data Sources Used to Construct Preliminary Database.

To address our study questions and test our hypothesis, we begin with a list of firms in the 42nd to 46th editions of the Corporate Giving Directory (CGD), which provides detailed information about

over 1,000 largest corporate foundations and programs in the U.S. The CGD offers a comprehensive list of major U.S. corporate giving programs, categorized by operations, location, and service areas, and has been utilized in prior research examining corporate community contributions in the U.S. market (Brammer 2005; Saiia et al. 2003).

The CSR data used in this study are constructed using two complementary AI-driven extraction frameworks that combine web-scale news processing with multi-source disclosure analysis. This section describes the construction process step by step and provides numerical detail on sample sizes, filtering, and validation at each stage.

CSR Data Construction

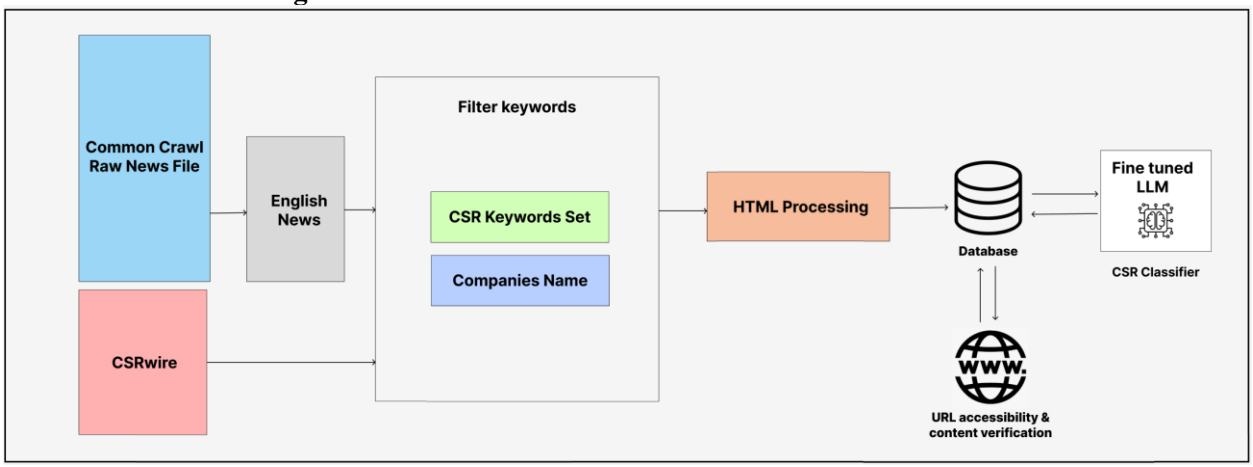
We begin with a universe of approximately 1.2 billion news articles published between 2000 and 2025 and archived in the Common Crawl News Archive and CSRwire. Because CSR-related disclosures constitute a small fraction of general news content, processing the full corpus directly would be computationally infeasible and introduce substantial noise. We therefore apply a high-recall screening procedure based on two conditions: (i) the presence of CSR-related keywords and (ii) the mention of at least one corporate donor identified by the CGD.

This initial screening step reduces the corpus from roughly 1.2 billion articles to approximately 116 million candidate articles, representing about 9.7% of the raw news universe. While this step is intentionally permissive by design, this stage eliminates articles that are clearly unrelated to corporate social activity, such as political commentary or general business news without CSR relevance.

Next, we use a fine-tuned Large Language Model (LLM) classifier to distinguish CSR-relevant articles from non-CSR content. The classifier is trained on a balanced labeled dataset consisting of 95,932 articles, split evenly between CSR-positive examples (primarily drawn from CSRWire)

and non-CSR news. When applied to the screened corpus, the classifier retains approximately 5.2 million articles, corresponding to 4.5% of the screened sample and 0.44% of the news universe. The resulting CSR-tagged articles are then subjected to a series of cleaning and validation steps. First, duplicate articles arising from press release syndication and reposting across multiple outlets are removed using normalized title matching. This step reduces the sample from 5.2 million to approximately 2.1 million unique articles. Second, we perform URL existence checks and content validation to ensure that each article remains accessible and that the stored text accurately reflects the original source. Articles failing these checks are discarded, yielding a final validated news corpus of approximately 1.1 million unique CSR-related articles. Figure 1 summarizes the news data extraction workflow.

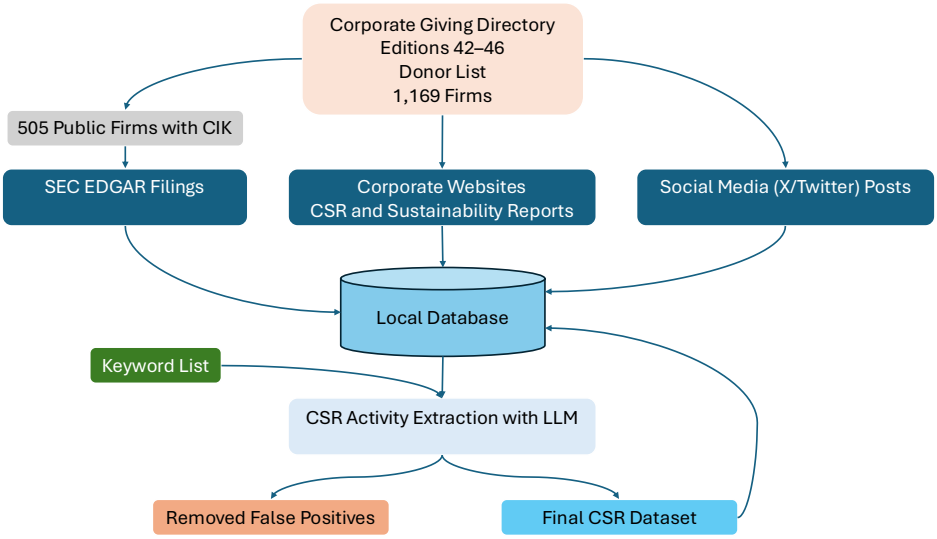
Figure 1: News-based CSR data construction workflow.



In parallel, we construct CSR data from regulatory filings, corporate sustainability reports, and social media disclosures. We begin with a cohort of 1,169 firms identified from editions 42 through 46 of the CGD. These firms serve as the canonical donor list for all extraction activities. From this cohort, we successfully match 505 publicly traded firms to Central Index Key (CIK) identifiers, enabling systematic retrieval of SEC filings. Using the SEC EDGAR system, we collect multiple filing types for each firm year, including 10-K, 10-Q, and 8-K filings, resulting in tens of thousands

of firm-year documents. In addition, we collect firm-published CSR and sustainability reports from corporate websites and historical social media posts from official firm accounts. All raw documents are stored in a structured database to facilitate batch processing and reproducibility.

Figure 2: Schematic representation of the multi-source data collection and LLM-driven CSR activity extraction pipeline.



Our initial sample period spans the years 2000-2025. The CGD indicates that sparse donation data were available from 2000-2015, but we used advanced technology to extend this sample to 2025. Iterative refinement remains ongoing, with keyword tuning and proxy updates planned as additional CSR disclosures emerge, embodying Salamon et al.'s (2011) view of valuation as evolving rather than definitive.

Methodology for Quantifying Nonmonetary Corporate Social Responsibility Donations

To extract CSR activity from these sources, we apply a keyword-anchored context window strategy. Rather than processing entire documents—which frequently exceed 50,000 words—we extract focused text windows around CSR-related terms. These windows are analyzed using an LLM instructed to record CSR activity only when an explicit, voluntary transfer of value is present and the recipient qualifies as a social or community organization. Each extracted record is then passed through a second validation stage that removes false positives, such as employee

compensation, commercial sponsorships, regulatory compliance costs, litigation settlements, and political spending. As shown in the final stages of the pipeline in Figure 2, this dual-stage extraction and validation process substantially reduces measurement error and ensures that retained observations reflect genuine CSR activity. Figure 2 summarizes the schematic representation of the multi-source data collection and LLM-driven CSR activity extraction pipeline.

The quantification of non-monetary CSR activity presents significant methodological challenges, as these activities are typically reported in descriptive rather than monetary terms. This study develops a systematic approach grounded in established nonprofit valuation frameworks, specifically drawing from Salamon et al. (2011) who emphasize the economic significance of non-market contributions and advocate for replacement-cost and opportunity-cost methods to render them comparable to cash donations. The methodology proceeds through four sequential stages: conceptual classification, conversion into units, shadow pricing, and uncertainty assessment, ensuring both transparency and replicability as recommended by Brown (1999) in her foundational work on volunteer activity valuation.

Each CSR activity can be conceptually classified into one of six mutually informative categories: fee relief, employee volunteering, in-kind products, training/capacity building, infrastructure access, and direct monetary donations. Classification employs keyword matching against the CSR activity description field, with terms such as “fee,” “ATM,” and “refund” indicating fee relief programs where companies waive transaction costs for community members; “volunteer,” “employee,” and “time” identifying labor contributions; “product,” “donation,” “mask,” and “equipment” flagging physical goods transfers; “training,” “seminar,” “workshop,” and “education” denoting knowledge transfer activities; “land,” “facility,” “infrastructure,” and “build”

signaling property or facility access; and direct monetary values exceeding $500 serving as the highest-confidence category. This taxonomy aligns with Salamon et al.'s (2011) call for clear conceptual boundaries in non-monetary CSR activity and Fennell et al.'s (2020) observation that quantifier type influences interpretability. Subsequent valuation is tailored to activity while binary indicator variables (0/1) enable flexible aggregation and overlap analysis.

Narrative descriptions are then converted into standardized countable units — a step Brown (1999) identifies as essential for aggregation across heterogeneous organizations. For employee volunteering and training, descriptive claims are standardized to full-time equivalent (FTE) workers, with vague references defaulting to 20 FTEs representing substantial but conservative contributions. In-kind products assume a fair market value of $150 per unit, scaled by employee count proxies to estimate quantity when unspecified. Fee relief programs are unitized as 10,000 waived transactions at $50 each, adjusted by geographic scope multipliers (local=1x, regional=2x, national=5x, international=10x) per Salamon et al.'s (2011) scope analysis. Infrastructure contributions employ rental equivalents of 100 FTE wages scaled by scope, while direct-reported amounts use the stated figure verbatim. These proxies, while assumption-laden, follow the ILO Manual on the Measurement of Volunteer Work (2011) in normalizing time-based contributions and Brown's (1999) standardization imperative to enable cross-organizational comparability.

Shadow pricing assigns economic values to unitized contributions using replacement-cost methodology, where the cost to procure equivalent market services substitutes for non-market donations (Salamon et al., 2011). Baseline wages derive from ILO Global Wage Reports (2000-2025), providing annual benchmarks ($5,000 in 2000 rising to $11,100 in 2025) adjusted by industry SIC codes (e.g., 1.15x multiplier for finance, 0.90x for retail) to reflect sector wage premiums. Volunteering and training are thus valued at hourly rates multiplied by estimated hours

and participants; products employ fair market equivalents; fee relief use transaction replacement costs; and infrastructure leverages rental proxies. This approach yields a total estimated value of $6 billion across 14,990 donations, with monetary donations (4.8% of records) anchoring high-confidence subsets while imputed values flag assumptions explicitly. The method's transparency—documenting unit conversions, wage sources, and multipliers—addresses Salamon et al.'s (2011) warning against false precision in non-cash valuation.

Uncertainty assessment operationalizes Fennell et al.'s (2020) findings on quantifier ambiguity through a confidence scoring system (1-10 scale) integrated as both row-level scores and category defaults. Monetary donations receive 9/10 due to explicit quantification; fee relief and volunteering score 7-8/10; products and descriptive infrastructure receive 6-7/10 reflecting market assumption sensitivity; and uncategorized defaults score 5/10. Scores compound data completeness (presence of year, SIC, employee count) and description specificity, yielding dataset average of 5.67/10. Binary indicators facilitate robustness checks, such as comparing direct-reported totals against imputed volunteering values, while overlap analysis (11% of donations match multiple categories) reveals classification nuance without double-counting in aggregates. This scoring aligns with Brown's (1999) emphasis on contestable assumptions in volunteer valuation and the ILO Manual's (2011) call for uncertainty disclosure.

Table 1 shows the initial universe consisting of 299,558 CSR donation observations spanning 1,072 firms over the 2000–2025 period. Many observations are dropped due to missing Compustat data, reducing the sample by 181,314 donation observations and 803 firms. A further screen requiring the availability of Institutional Shareholder Service (ISS) data removes an additional 26,849 donation observations and 105 firms. These steps alone underscore the importance of reliable firm-level financial reporting for CSR research. The regressions that will be presented in

the next section rely on the preliminary database's sample of 91,395 firm-donation-year observations across 164 firms.

After compiling a comprehensive set of corporate donations, firm and governance data, we address our study questions and test our hypothesis. Equation 1 is designed to test our initial hypothesis. In this model, we regress the logarithm of total donations (LTOTALDON) on determinants of corporate donations as well as the log proximity (LPROXIMITY) of recipients to firms. The CGD provides us with names of recipients, making it feasible to identify their location.

$$ltotaldon_{it} = f\,(lproximity_{it}, corporate\ governance_{it}, Firm\ Controls_{it}, Industry_{it}, Year_{it}) \quad (1)$$

Following prior research (Ding et al., 2019; Benabou and Tirole, 2010; Ali et al., 2017; Jo and Harjoto 2012; Udayasankar 2008), we add a set of control variables that explain the variations in ltotaldon, such as corporate governance, firm size, industry, and profitability. We also control for Fama and French industry and year-fixed effect. Tables 5 and 6 show the coefficients from that regression analysis. The coefficients indicate that proximity is an important determinant and operates differently across CSR categories, reinforcing the need to model CSR as a heterogeneous set of strategic choices.

The magnitude of the attrition shown in Table 1 is economically meaningful rather than mechanical. While roughly 30 percent of the original donation observations survive all data requirements, these retained observations represent firms with systematically richer disclosure environments and governance data that can be used for regressions and subsequent analyses. This selection aligns with the study's objective of linking CSR behavior to firm characteristics, governance, and strategic considerations. However, the final database will also include CSR by companies without such detailed data so that broad measures of CSR activity can be calculated. These broad measures of CSR activity will advance economics by systematically monetizing

otherwise untracked activity, bridging narrative CSR reports to empirical analysis while maintaining auditability through documented assumptions, sources, and confidence metrics.

[Insert Table 1 here]

IV. RESULTS

In addition to building a CSR database that would enable users to calculate its economic impact, the project addresses several pivotal research questions:

1) Have decisions associated with CSR evolved over time?

2) Which types of firms engage in CSR for their local community?

3) Why do firms initiate CSR activities?

The analysis draws on a preliminary AI-constructed CSR dataset spanning 2000-2025, with 91,395 firm-donation-year observations across 164 firms. This study merges the AI-constructed dataset with control variables such as corporate governance, firm size, industry, profitability, and socio-economic conditions of the firm's geographic headquarters location. Industry fixed effects (using 48 Fama and French industry codes) and year fixed effects are also included. The CSR results discussed in this section and shown in Appendix C are all calculated after controlling for the standard control variables.

Have decisions associated with CSR evolved over time?

Figure 3 in Appendix C presents a stacked bar chart tracking the evolution of total CSR donations from 2000 to 2025, disaggregated into Total, Nonmonetary, and Monetary components in millions of dollars. The chart shows a clear upward trajectory in aggregate giving, with total donations growing consistently from under $1 billion around 2000 to peaks exceeding $2.5 billion by 2025. Monetary donations dominate throughout the sample period, comprising 60-70% of totals—starting at roughly $0.6 billion in 2000 and reaching over $1.7 billion by 2025—reflecting their

ease of quantification and reporting. Non-monetary CSR activity, the second-largest category, begins at about 20 percent in the early 2000s but accelerates post-2018, surging to 35-40 percent during 2020-2022 peaks, likely due to COVID-19-driven volunteerism and in-kind medical supplies. This non-monetary CSR activity is valued via ILO wages and replacement costs.

[Insert Figure 3 here]

Figure 4 in Appendix C consists of two panels comparing donation trends for firms with and without corporate foundations across 2000-2025. The "With Foundation" panel displays dramatically higher levels, consistently 2-3 times greater—ranging from $0.8-1.2 billion annually in early years to over $2 billion by 2025—versus non-foundation peers at $0.3-0.6 billion early on and $0.7-1 billion later. Foundation firms exhibit remarkable stability, with minimal dips during the 2008-2009 crisis (dropping only ~10-15 percent), while non-foundation giving declines more sharply (~30-40 percent). The gap underscores foundations as institutionalized commitments, creating insulated funding for sustained investment beyond profit cycles. Higher non-monetary proportions in foundation totals (visually ~30-40 percent of stacks) reflect structured programs like employee volunteering, contrasting ad-hoc responses and affirming strategic CSR evolution.

[Insert Figure 4 here]

Which Types of Firms Engage in CSR for their Local Community?

Figure 5 in Appendix C illustrates industry-wide donation shares, with pharmaceuticals leading at 25.90%, followed by telecommunications (23.39%), utilities (22.85%), electrical equipment (10.13%), and transportation (3.45%), summing across 48 sectors. Finance contributes ~25% share ($1.5 billion total), while others like retail trail lower. Absolute values scale with overall growth, from smaller pre-2010 bases to dominant post-2020 contributions.

Prior research (Choi et al. 2019) sheds light on the importance of donation appeals and geographic proximity (local versus global). There is a tendency to believe that donating to local communities is more impactful than donating to global causes (Grau and Folse 2007). However, certain areas have disadvantages and require more charitable attention than others. It is interesting to observe U.S. corporations donating to global organizations, and this phenomenon requires further investigation. Relationships between firm characteristics and proximity are shown in Table 5 of Appendix B. To summarize, large, foundation-backed firms in regulated or litigated industries are most heavily engaged in local CSR. Conversely, internationally oriented firms in pharma and tech sectors lead global CSR.

Figure 5 in Appendix C reveals sector-specific propensities for local CSR. Key findings highlight distinct firm profiles in CSR engagement: regulated industries like utilities and telecom prioritize local communities through stable foundation-backed giving, while global sectors such as pharmaceuticals extend to broader causes, driven by strategic factors like firm size, complexity, mergers, market power, and governance traits including board diversity. Proximity consistently predicts donation intensity, rejecting the null hypothesis *H0* of random allocation and informing heterogeneity in CSR drivers and enabling targeted policy.

[Insert Figure 5 here]

Figures 6 and 7 in Appendix C track geographic proximity and scope. Figure 6 shows average geographic proximity (log distance from headquarters to recipient) annually from 2000-2025, showing a gradual decline from ~4.5-5.0 log miles early on to ~3.8-4.2 by 2025, indicating tightening focus on nearer causes. Dips occur around crises (2008: ~4.3; 2020: ~3.9), with post-2018 stabilization reflecting institutionalized localism amid rising totals. In contrast, Figure 7 breaks CSR activity by proximity category, with national bars tallest (~70-80% share, $1.5-$2

billion peaks), local mid (~15-20%, $0.4-0.6 billion), and international thin but growing (~$0.05-0.1 billion by 2025). The trends shown in Figures 6 and 7 reject random allocation and support stakeholder theory. In other words, firms have increasingly prioritized local CSR, with implications for regional spillovers.

[Insert Figure 6 & 7 here]

Why Do Firms Initiate CSR Activities?

Table 2 reports the industry distribution of donations based on the Fama–French 48 industry classification. The distribution is highly skewed, with a small number of industries accounting for the majority of observed CSR activity. Pharmaceuticals represent the largest single contributor, with 23,671 donation observations, accounting for 25.90 percent of total donations in the final sample. This dominance is economically intuitive given the sector's global footprint, high margins, and strong incentives to engage in health-related and science-oriented CSR initiatives. Telecommunications and utilities follow closely, contributing 23.39 percent and 22.85 percent of donations, respectively. Together, these three industries alone account for more than 72 percent of all observed donations. Beyond these dominant sectors, electrical equipment firms contribute 10.13 percent of donations, while transportation accounts for 3.45 percent. Retail firms contribute a nontrivial 4.63 percent, reflecting their broad consumer exposure and community engagement incentives. In contrast, many industries contribute negligibly, with several sectors such as rubber and plastics, stone, and entertainment each accounting for less than 0.05 percent of donations. This sharp concentration highlights substantial heterogeneity in CSR engagement across industries and motivates the inclusion of industry fixed effects and industry-specific analyses in later tables.

Table 2 establishes an important foundation for empirical analysis. CSR activity is far from evenly distributed across industries, with pharmaceuticals emerging as the single most important

contributor in terms of donation frequency, followed closely by telecom and utilities. The industry composition foreshadows the later regression results, where globally oriented and regulated industries exhibit systematically different CSR patterns in terms of proximity, scope, and magnitude. Table 1 thus serves not only as a descriptive benchmark but also as a roadmap for interpreting the cross-sectional and multivariate results that follow.

[Insert Table 2 here]

Table 3 presents descriptive statistics for the main variables used in the empirical analysis and provides important context for interpreting the regression results that follow. The mean value of LTOTALDON is 6.851, corresponding to approximately $950 million in total annual donations when converted from logs, with a standard deviation of 2.511, indicating cross-firm and intertemporal variation in CSR intensity. Monetary donations dominate in levels, with LMON averaging 6.524, while nonmonetary donations remain economically meaningful, as reflected by the distribution of LNONMON. The wide gap between the minimum and maximum values of LTOTALDON, ranging from 0.531 to 15.257, suggests that the sample captures both relatively small, localized donation activity and extremely large, institutionalized CSR programs.

Geographic proximity exhibits meaningful dispersion. LPROXIMITY has a mean consistent with donations occurring primarily at national rather than purely local distances, yet the interquartile range indicates that a substantial fraction of donations are geographically close to firm headquarters. This variation is central to the paper's identification strategy. Approximately 75.5 percent of donations are national in scope, while only about 1.0 percent are international, confirming that global CSR, though limited in frequency, may still exert outsized effects in magnitude. Firm characteristics also display strong heterogeneity: LNAT (log total assets) indicates a sample dominated by large firms, while COMPLEX has a mean near one, reflecting

that the vast majority of firms report foreign or extraordinary activities. Governance variables, such as board size and gender composition, show meaningful variation that later helps explain differences in CSR intensity.

[Insert Table 3 here]

Table 4 reports pairwise correlations among the main variables used in the analysis. Total donations (LTOTALDON) are strongly and positively correlated with monetary donations (LMON) ($\rho = 0.85$, $p < 0.01$) and positively correlated with nonmonetary donations (LNONMON) ($\rho = 0.26$, $p < 0.01$), indicating that both components contribute to overall giving. Monetary and nonmonetary donations are negatively correlated with one another ($\rho = -0.37$, $p < 0.01$), suggesting that firms tend to emphasize one form of giving over the other rather than increasing both simultaneously. Geographic proximity (LPROXIMITY) is positively correlated with total donations ($\rho = 0.08$, $p < 0.01$), and non-monetary donations ($\rho = 0.14$, $p < 0.01$), but negatively correlated with monetary donations ($\rho = -0.04$, $p < 0.01$). Proximity is negatively correlated with national scope ($\rho = -0.31$, $p < 0.01$), national operations ($\rho = -0.24$, $p < 0.01$), and international operations ($\rho = -0.18$, $p < 0.01$), reflecting expected mechanical relationships between distance and donation scope. Firm characteristics exhibit generally modest correlations with donation measures. Market concentration (HHI) is positively correlated with total donations ($\rho = 0.14$, $p < 0.01$), while mergers and acquisitions activity (MA) is weakly positively correlated with total donations ($\rho = 0.00$, $p < 0.01$). Firm size (LNAT) is negatively correlated with total donations ($\rho = -0.31$, $p < 0.01$), consistent with larger firms donating across broader geographic scopes. Governance variables show limited bivariate associations: the number of directors is positively correlated with total donations ($\rho = 0.12$, $p < 0.01$), whereas the percentage of female directors is negatively correlated with total donations ($\rho = -0.23$, $p < 0.01$). Overall, the correlations are generally small to moderate

in magnitude, and no pairwise correlations—outside of mechanically related donation components—suggest severe multicollinearity concerns (using cutoff 0.8 correlation), supporting the use of multivariate regression analysis.

[Insert Table 4 here]

Corporate Donations and Geographic Proximity

Table 5 reports the multivariate regressions examining the determinants of charitable donations, with LTOTALDON as the dependent variable. Across all three models, geographic proximity (LPROXIMITY) is a strong and economically meaningful predictor of donation intensity. In Model (1), the coefficient on LPROXIMITY is −0.0561 with a t-statistic of −15.21, indicating that a 1 percent increase in distance between the firm and the recipient is associated with approximately a 5.6 percent decrease in total donations. This effect remains highly significant in Models (2) and (3), though attenuated in magnitude once additional controls and governance variables are introduced, with coefficients of −0.0196 and −0.0185, respectively. The stability of the sign and significance across specifications provides strong evidence against random geographic allocation of CSR. The results in this section reject the null hypothesis and support the alternative that there is an association between the geographic proximity of donations and total donations.

Several firm-level characteristics exhibit large economic effects. The coefficient of NONMON is positive and highly significant, exceeding 1.11, implying that non-monetary donations are more than three times larger in total value than monetary-only donations, consistent with the replacement-cost valuation approach. BIGN, indicating Big 4 audit status, has a coefficient of approximately 3.30, suggesting that firms with higher reporting quality and visibility engage in substantially greater CSR activity. Market concentration, measured by HHI, is also positive and significant, with coefficients between 0.77 and 0.85, indicating that firms with greater market

power donate more. Mergers and acquisitions activity has one of the largest effects, with coefficients around 1.45, implying nearly a 145 percent increase in donations during M&A years, consistent with reputational and stakeholder management motives.

Governance and risk variables further reinforce the strategic interpretation of CSR. REST and MW are both positive and highly significant, indicating that firms facing reporting issues or internal control weaknesses donate more, potentially mitigate reputational risk. LNAT is positive and significant, confirming that larger firms donate more in absolute terms. COMPLEX has a coefficient exceeding 1.60, underscoring the role of organizational scope in shaping CSR activity. NATIONAL donations are significantly smaller than local ones, with coefficients around −0.35, while INTERNATIONAL donations are significantly larger, with coefficients around 0.29–0.33, despite their rarity. The inclusion of board characteristics in Model (3) reveals that larger boards and a higher percentage of female directors are associated with lower total donations, with PCT_FEMALE exhibiting a particularly large negative coefficient of −1.975, suggesting a reallocation rather than expansion of CSR budgets.

[Insert Table 5 here]

Table 6 extends the analysis by exploring cross-sectional heterogeneity and category-specific effects. Panel A focuses on firms with corporate foundations and shows that proximity effects are substantially amplified in this subsample. The interaction between LPROXIMITY and FOUNDATION is −2.5863 with a t-statistic of −18.16, indicating that foundation-backed firms are far more likely to concentrate on donations locally. NATIONAL remains negative and highly significant, while INTERNATIONAL is positive, reinforcing the notion that foundations facilitate both deep local engagement and selective global outreach. The explanatory power of these models

increases modestly, with R-squared values exceeding 48 percent, highlighting the importance of institutional CSR structures.

Panel B of Table 6 examines donation behavior by category and reveals sharp heterogeneity in proximity sensitivity. Education, religion, civic and public affairs, health, and social services all exhibit positive and significant interactions with LPROXIMITY, indicating that these categories are disproportionately local. For example, the interaction coefficient for health is approximately 0.136, suggesting that health-related donations attenuate the negative distance effect and are less sensitive to geographic separation. In contrast, science-related donations show no statistically significant interaction with proximity, consistent with their global orientation. Arts and humanities display a strongly negative interaction, indicating that these donations are highly localized.

The baseline coefficient on LPROXIMITY in Panel B is negative and statistically significant across all specifications, ranging from −0.0968 to −0.0439 ($p < 0.01$), confirming that, on average, charitable donations decline as geographic distance between the firm and recipient increases. However, the interaction terms reveal pronounced heterogeneity across donation categories.

Education-related donations exhibit positive and statistically significant interactions with proximity, with coefficients ranging from 0.0397 to 0.0452 and significance at the 5 percent level or better. This indicates that distance meaningfully attenuates the baseline decline in giving for education-related causes. Religion and civic and public affairs display even stronger local orientation. The interaction coefficients for religion range from 0.0962 to 0.0994 and are significant at the 1 percent level, while civic and public affairs interactions range from 0.0737 to 0.0970, also significant at the 1 percent level. These magnitudes imply that donations in these categories are substantially concentrated in geographically proximate communities than the average CSR dollar.

Health-related donations exhibit the largest and most robust interaction effects. The coefficients on LPROXIMITY × HEALTH range from 0.0966 to 0.1356 and are highly significant at the 1 percent level across all models. Economically, these estimates imply that the negative association between distance and giving is sharply reduced for health-related causes, consistent with hospitals, clinics, and community health organizations being tightly embedded in local service areas.

Social services donations also show positive and statistically significant proximity interactions, with coefficients ranging from 0.0383 to 0.0646 and significance at the 5 percent level or better. Together, these results indicate that categories tied to direct community provision—education, religion, civic engagement, health, and social services—are disproportionately local in nature, even after controlling for firm characteristics, governance, and donation scope.

In contrast, science-related donations show no statistically significant interaction with proximity in any specification, with coefficients that are small in magnitude and statistically indistinguishable from zero. This suggests that scientific and research-oriented CSR is geographically neutral or globally oriented, consistent with beneficiaries such as universities, research institutes, and laboratories operating within national or international knowledge networks where physical distance is less salient.

Arts and humanities donations stand out in the opposite direction. The interaction between proximity and arts and humanities is negative and highly significant across all models, with coefficients clustered around -0.014 ($p < 0.01$). This result implies that arts-related giving becomes more sensitive to distance than the baseline, reinforcing the idea that cultural and artistic donations are strongly tied to immediate local communities and lose relevance rapidly as geographic separation increases.

Environmental donations exhibit weaker and less stable interaction effects. While the interaction term is positive and marginally significant in the simplest specification (coefficient of 0.0822, $p < 0.10$), it becomes statistically insignificant once additional controls are introduced. This pattern suggests that environmental CSR occupies an intermediate position between local and global orientation, reflecting the coexistence of place-based environmental initiatives and contributions aimed at broader public goods.

Taken together, Panel B demonstrates that geographic proximity is not a uniform determinant of CSR activity. Instead, proximity interacts systematically with the underlying purpose of the donation. Categories associated with direct community services exhibit strong local concentration, while knowledge-based and globally framed causes are largely insensitive to distance. These results reinforce the core findings of Table 5 while extending them by showing that proximity effects operate through category-specific channels rather than through aggregate CSR alone.

These results demonstrate that proximity operates differently across CSR categories, reinforcing the need to model CSR as a heterogeneous set of strategic choices rather than a single aggregate outcome.

[Insert Table 6 here]

V. CONCLUSION AND RECOMMENDATIONS

This study develops a scalable, AI-driven framework that fundamentally transforms how corporate social responsibility (CSR) is measured, validated, and analyzed. By integrating NLP with LLM validation across diverse and previously fragmented sources, the study constructs a unified and economically interpretable CSR database spanning 2000–2025. In doing so, it addresses one of the most persistent limitations in CSR research—namely, the lack of consistent, comparable, and comprehensive data.

The empirical findings underscore several important patterns. First, CSR activity has grown substantially over time, both in scale and in composition, with non-monetary contributions becoming increasingly significant after 2018. Second, CSR engagement is highly concentrated across industries, with pharmaceuticals, telecommunications, utilities, and finance accounting for the majority of activity. Third, and most importantly, the analysis provides strong evidence that CSR allocation is not random: geographic proximity plays a central and economically meaningful role, particularly for community-oriented categories such as health, education, and social services. These findings collectively reinforce the view that CSR is a strategic and heterogeneous firm decision shaped by governance, market structure, and stakeholder considerations rather than a purely altruistic or symbolic activity.

The implications of these findings are substantial for multiple stakeholders. For researchers, the database offers a long-awaited opportunity to move beyond coarse proxies and subjective ratings toward causal identification of CSR's economic and social effects. The ability to link granular CSR actions to firm-level financial data, governance structures, and geographic outcomes opens new avenues for understanding how CSR influences firm value, risk, labor markets, and regional development.

For investors and financial analysts, the framework enhances transparency and comparability in CSR reporting. By monetizing non-monetary contributions using replacement-cost methods and introducing confidence metrics, the database reduces the risk of greenwashing and allows for more rigorous integration of CSR into valuation models, portfolio allocation, and ESG screening. This is particularly important given the documented tendency of firms to substitute between monetary and non-monetary CSR activities, which may otherwise obscure the true scale of corporate contributions.

For corporate managers, the findings highlight the strategic nature of CSR. The strong association between CSR and factors such as market power, mergers and acquisitions, governance quality, and reputational risk suggests that CSR is actively used as a tool for stakeholder management and risk mitigation. The documented heterogeneity across industries and donation categories further implies that firms can optimize CSR strategies by aligning them with their operational footprint, stakeholder base, and long-term positioning.

For policymakers and standard-setting bodies, the study provides a concrete foundation for the development of standardized CSR reporting frameworks. Much like financial accounting evolved toward generally accepted principles, the structured and validated nature of this database demonstrates the feasibility of establishing consistent measurement standards for CSR. Such standards could reduce information asymmetry, improve regulatory oversight, and enhance the credibility of corporate disclosures. Moreover, the integration of non-monetary CSR into economic measurement has important implications for national accounting. Incorporating these activities into official statistics could lead to more accurate estimates of economic output and societal welfare, particularly if real-time data pipelines are developed to support timely policy decisions.

Looking ahead, several directions for future research emerge. A primary avenue is causal analysis: future studies can exploit the database to identify the impact of CSR on firm performance, cost of capital, labor outcomes, and regional economic development using quasi-experimental designs. Improvements in LLM accuracy, entity resolution, and real-time data ingestion could further reduce measurement error and enable dynamic tracking of CSR activity at higher frequency. Therefore, future research could incorporate real-time data ingestion pipelines, potentially leveraging agentic AI systems capable of continuously monitoring, extracting, and validating CSR disclosures as they emerge.

Additionally, future work could extend the framework internationally, allowing for cross-country comparisons of CSR behavior and its interaction with institutional environments, regulatory regimes, and cultural norms. This would provide valuable insights into how global firms allocate CSR resources across jurisdictions and how policy differences shape corporate behavior.

Further research should also explore the interaction between CSR and other dimensions of ESG, particularly environmental and governance factors, to better understand complementarities and trade-offs. The observed substitution between monetary and non-monetary contributions suggests that firms optimize across multiple CSR channels, raising important questions about efficiency, effectiveness, and stakeholder impact.

Finally, the integration of CSR data into macroeconomic systems presents a promising frontier. Future work could develop methodologies for incorporating CSR into satellite accounts or adjusted GDP measures, capturing the economic value of activities that are currently excluded from official statistics. Real-time CSR tracking could also support policy responses during crises, as evidenced by the surge in non-monetary contributions during the COVID-19 period.

In sum, this study not only provides a novel measurement framework but also establishes a foundation for a new generation of CSR research and policy analysis. By transforming fragmented disclosures into structured, comparable, and economically meaningful data, it bridges the gap between corporate behavior and societal outcomes, enabling more informed decisions across academia, industry, and government.

Appendix A: An example of 1st Source Corporation's Donations scattered across social media, and public filings

Date	Donations	Source
05/22/2020	Donated financial advice and provided payment relief to 550 businesses worth hundreds of millions of dollars.	Press release
August 8, 2023	Donated $27,000 grant to the La Porte County Family YMCA	News Media
2023	12,000 volunteer hours, 500 organizations helped, 312 employee volunteers, 3,000,000 donations	Foundation website
2021	Donated more than $3 million to the communities, supporting education, social services, the arts, economic development, and community organizations. Volunteered more than 15,000 hours to community organizations.	Foundation website
October 11, 2021	Various community organizations in St. Joseph County and Berrien County, Michigan have received a total grant of $120,000 grant from 1st Source Corp.	News Media
April 25, 2018	Donated $14,200 to Reins of Life.	News Media
August 25, 2020	Donated $30,000 in support of frontline workers.	News Media
December 31, 2022	We reimbursed over $137,000 to colleagues for tuition reimbursement with an average of $3,600 per colleague who used the benefit…our colleagues donated approximately 13,000 hours to a total of 900 different organizations…our colleagues contributed over $189,000 to local United Way organizations…contributed over $745,000 to over 450 deserving and successful community service organizations.	Annual report
March 10, 2023	In 2022, we gave 191 formal presentations and reached over 2,500 employees of our clients and community members. We also provide financial education and information on our online and mobile banking platform.	Proxy Statement
September 12, 2023	Donated approximately $393,343, raised through corporate and customer donations.	News Media

Appendix B: Variable Definitions

Variable	Definition
LTOTALDON	Natural logarithm of total monetary and nonmonetary donations
LPROXIMITY	Natural logarithm of the distance in miles between the donor and the recipient
EDUCATION	An indicator variable equal to 1 if the donations were made for education, zero otherwise
RELIGION	An indicator variable equal to 1 if the donations were made for religion, zero otherwise
CIVICPUBLICAFFAIRS	An indicator variable equal to 1 if the donations were made for civic and public affairs, zero otherwise
SCIENCE	An indicator variable equal to 1 if the donations were made for science, zero otherwise
HEALTH	An indicator variable equal to 1 if the donations were made for health, zero otherwise
ARTS_HUMANITIES	An indicator variable equal to 1 if the donations were made for arts and humanities, zero otherwise
ENVIRONMENT	An indicator variable equal to 1 if the donations were made for the environment, zero otherwise
SOCIALSERVICES	An indicator variable equal to 1 if the donations were made for social services, zero otherwise
NONMON	An indicator variable equal to 1 if the donations were nonmonetary, zero otherwise
FOUNDATION	An indicator variable equal to 1 if the donor has a foundation, zero otherwise
BIGN	An indicator variable equal to 1 if the donor's financial statements are audited by a Big 4 auditor (KPMG, EY, Deloitte, or PwC), zero otherwise
HHI	The Herfindahl–Hirschman Index (HHI), calculated as the sum of the squared market shares of all firms in a market, where each firm's market share is expressed as a proportion of total market activity
MA	An indicator variable equal to 1 if the donor engaged in mergers and acquisitions, zero otherwise
SGROWTH	Sales growth measured as the change in sales from year t−1 to year t
ROA	Net income divided by lagged total assets
Z	Altman's Z-score
REST	An indicator variable equal to 1 if the donor's financial statements contain restatements, zero otherwise
MW	An indicator variable equal to 1 if the donor's financial statements contain material weaknesses, zero otherwise
LNAT	Natural logarithm of the donor's total assets
COMPLEX	An indicator variable equal to 1 if the donor reported foreign or extraordinary activities, zero otherwise
LIT	An indicator variable equal to 1 if a firm operates in a litigation-intensive industry based on its SIC code (2833–2836, 8731–8734, 7371–7379, 3570–3577, or 3600–3674), zero otherwise
NATIONAL	An indicator variable equal to 1 if the donations were made for a national location, zero otherwise
INTERNATIONAL	An indicator variable equal to 1 if the donations were made for an international location, zero otherwise
N_DIRECTORS	Number of members on the donor's board of directors
PCT_FEMALE	Percentage of female directors on the donor's board

Table 1: Sample Statistics

Panel A: Sample Selection

	No. Donations	*Firms*
CSR Charitable Donations (2000-2025)	299,558	1072
(-) Firms with missing data from Compustat	(181,314)	(803)
	118,244	269
(-) Firms with missing data from ISS ESG*	(26,849)	(105)
= Final Sample	**91,395**	**164**

* ISS ESG data refers to environmental, social, and governance (ESG) data and ratings produced by Institutional Shareholder Services (ISS) — one of the major global providers used in academic research, asset management, and corporate governance analysis.

Table 2: Sample Distribution by Industry

Industry Name	FF48	Donations	%Donation
Agriculture	1	20	0.02%
Oil	3	1661	1.82%
Stone	4	8	0.01%
Construction	5	38	0.04%
Food	6	1964	2.15%
Tobacco	7	13	0.01%
Textiles	8	269	0.29%
Lumber	9	5	0.01%
Furniture	10	704	0.77%
Paper	11	161	0.18%
Printing	12	47	0.05%
Chemicals	13	60	0.07%
Pharmaceuticals	14	23671	25.90%
Chemicals	15	524	0.57%
Petroleum	16	149	0.16%
Rubber & Plastics	17	3	0.00%
Stone, Clay, Glass	19	13	0.01%
Wholesale	20	533	0.58%
Machinery	21	1106	1.21%
Electrical Equipment	22	9255	10.13%
Automobiles	23	583	0.64%
Instruments	24	111	0.12%
Misc Manufacturing	25	79	0.09%
Transportation	26	3157	3.45%
Telecom	27	21377	23.39%
Utilities	28	20886	22.85%
Waste	29	39	0.04%
Wholesale	30	96	0.11%
Retail	31	4229	4.63%
Business Services	34	564	0.62%
Rubber & Plastics	36	3	0.00%
Entertainment	38	4	0.00%
Health	40	29	0.03%
Engineering	46	34	0.04%
		91395	100.00%

Table 3: Descriptive Statistics

Variable	Mean	Std Dev	Minimum	Maximum	25th Pctl	50th Pctl	75th Pctl
LTOTALDON	6.851	2.511	0.531	15.257	4.828	6.397	8.517
LMON	6.524	2.611	0.000	15.202	4.615	6.217	8.058
LNONMON	0.368	2.040	0.000	12.332	0.000	0.000	0.000
LPROXIMITY	5.415	1.917	0.000	9.291	4.304	5.412	7.046
NONMON	0.032	0.177	0.000	1.000	0.000	0.000	0.000
BIGN	0.988	0.110	0.000	1.000	1.000	1.000	1.000
HHI	0.114	0.130	0.033	0.953	0.058	0.071	0.130
MA	0.164	0.371	0.000	1.000	0.000	0.000	0.000
SGROWTH	0.088	0.195	-0.749	2.818	0.005	0.023	0.092
ROA	0.074	0.095	-0.258	0.414	0.029	0.058	0.074
Z	-0.746	0.937	-4.634	18.422	-0.738	-0.345	-0.205
REST	0.235	0.424	0.000	1.000	0.000	0.000	0.000
MW	0.000	0.019	0.000	1.000	0.000	0.000	0.000
LNAT	11.205	1.275	5.660	12.929	10.463	11.602	11.887
COMPLEX	0.956	0.205	0.000	1.000	1.000	1.000	1.000
LIT	0.366	0.482	0.000	1.000	0.000	0.000	1.000
NATIONAL	0.755	0.430	0.000	1.000	1.000	1.000	1.000
INTERNATIONAL	0.010	0.101	0.000	1.000	0.000	0.000	0.000
N_DIRECTORS	11.341	1.424	6.000	20.000	10.000	11.000	12.000
PCT_FEMALE	0.300	0.082	0.000	0.667	0.250	0.300	0.364

Table 4: Correlation Matrices

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1		0.85a	0.26a	0.08a	0.25a	0.04a	0.14a	0.00	0.03a	0.06a	-0.01	0.09a	0.03a	-0.31a	-0.18a	-0.18a	-0.24a	0.09a	0.12a	-0.23a
2	0.71a		-0.37a	-0.04a	-0.37a	0.03a	0.10a	-0.01	-0.03a	0.03a	-0.01	0.09a	0.01	-0.25a	-0.15a	-0.14a	-0.16a	-0.02	0.06a	-0.18a
3	0.36a	-0.37a		0.14a	0.99a	0.02a	0.17a	0.06a	0.00	0.05a	-0.02	-0.03a	0.03a	-0.06a	-0.04a	-0.05a	-0.15a	0.20a	0.04a	-0.02c
4	0.06a	-0.04a	0.14a		0.14a	-0.04a	0.15a	-0.12a	-0.13a	0.13a	-0.04a	-0.04a	0.00	0.01	0.05b	-0.21a	0.45a	0.18a	0.18a	0.03
5	0.34a	-0.37a	0.99a	0.13a		0.02a	0.17a	0.05a	-0.03a	0.01	0.00	-0.03a	0.03a	-0.06a	-0.04a	-0.05a	-0.15a	0.20a	0.04a	-0.02c
6	0.04a	0.03a	0.02a	-0.04a	0.02a		-0.38a	0.05a	0.03a	0.02a	0.13a	-0.03a	0.00	0.38a	0.26a	0.08a	0.02	0.01	0.18a	0.11a
7	0.30a	0.10a	0.26a	0.15a	0.27a	-0.38a		0.09a	-0.02	0.00	0.00	-0.26a	0.04a	-0.36a	-0.18a	-0.23a	-0.08a	0.11a	-0.26a	-0.06a
8	0.03a	-0.01	0.06a	-0.12a	0.05a	0.05a	0.09a		-0.18a	-0.04a	0.00	-0.18a	-0.01	-0.06a	-0.13a	0.28a	0.11a	0.00	-0.31a	0.39a
9	-0.06a	-0.03a	-0.03a	-0.13a	-0.03a	0.03a	-0.02a	-0.18a		-0.34a	0.12a	-0.13a	-0.01	0.01	0.05b	0.35a	0.13a	-0.02	0.16a	0.04a
10	0.04a	0.03a	0.02a	0.13a	0.01	0.02a	0.00	-0.04a	-0.34a		-0.76a	-0.27a	-0.01	-0.33a	0.03a	0.28a	-0.10a	-0.01	-0.12a	0.08a
11	-0.01	-0.01	-0.02	-0.04a	0.00	0.13a	0.00	0.01	0.12a	-0.76a		0.21a	0.00	0.43a	-0.06a	-0.47a	0.09a	-0.07a	0.07a	-0.02
12	0.07a	0.09a	-0.03a	-0.04a	-0.03a	-0.03a	-0.26a	-0.18a	-0.13a	-0.27a	0.21a		0.00	-0.04a	0.09a	-0.42a	-0.24a	-0.05a	0.63a	-0.36a
13	0.03a	0.01	0.03a	0.00	0.03a	0.00	0.04a	-0.01	-0.01	-0.01	0.00	0.00		-0.03a	0.00	-0.01	-0.02	0.00	0.00	-0.01
14	-0.31a	-0.25a	-0.06a	0.01	-0.06a	0.38a	-0.36a	-0.06a	0.01	-0.33a	0.43a	-0.04a	-0.03a		0.23a	-0.12a	0.19a	-0.03a	0.08a	0.34a
15	-0.18a	-0.15a	-0.04a	0.05b	-0.04a	0.26a	-0.18a	-0.13a	0.05b	0.03a	-0.06a	0.09a	0.00	0.23a		0.16a	0.13a	0.01	0.10a	0.06a
16	-0.18a	-0.14a	-0.05a	-0.21a	-0.05a	0.08a	-0.23a	0.28a	0.35a	0.28a	-0.47a	-0.42a	-0.01	-0.12a	0.16a		0.21a	0.00	-0.09a	0.23a
17	-0.24a	-0.16a	-0.15a	0.45a	-0.15a	0.02	-0.08a	0.11a	0.13a	-0.10a	0.09a	-0.24a	-0.02	0.19a	0.13a	0.21a		-0.18a	-0.14a	0.19a
18	0.09a	-0.02	0.20a	0.18a	0.20a	0.01	0.10a	0.00	-0.02	-0.01	-0.08a	-0.05a	0.00	-0.03a	0.01	0.00	-0.18a		-0.04a	-0.01c
19	0.09a	0.06a	0.04a	0.18a	0.04a	0.18a	-0.26a	-0.31a	0.16a	-0.12a	0.07a	0.63a	0.00	0.08a	0.10a	-0.09a	-0.14a	-0.04a		-0.38a
20	-0.23a	-0.18a	-0.02c	0.03	-0.02c	0.11a	-0.06a	0.39a	0.04a	0.08a	-0.02	-0.36a	-0.01	0.34a	0.06a	0.23a	0.19a	-0.01c	-0.38a	

The upper diagonal reports Spearman correlation coefficients, and the lower diagonal reports Pearson correlation coefficients. Statistical significance is indicated by letter markers: “a” denotes significance at the 1 percent level, “b” at the 5 percent level, and “c” at the 10 percent level. The numbers in Table 4 correspond to variables used in the analysis as follows: (1) total donations (ltotaldon), (2) monetary donations (lmon), (3) nonmonetary donations (lnonmon), (4) proximity (lproximity), (5) an indicator for nonmonetary donations (Nonmon), (6) big 4 (BIGN), (7) market concentration (HHI), (8) mergers and acquisitions activity (MA), (9) sales growth (SGROWTH), (10) return on assets (ROA), (11) Z-score (Z), (12) restatements (REST), (13) material weaknesses (MW), (14) log of total assets (LNAT), (15) organizational complexity (COMPLEX), (16) litigation exposure (LIT), (17) donations to national causes (National), (18) donations to international causes (International), (19) number of directors (n_directors), and (20) percentage of female directors (pct_female).

Table 5: Regression of the Determinants of Charitable Donations

	Model (1) *DV=LTOTALDON*		Model (2) *DV=LTOTALDON*		Model (3) *DV=LTOTALDON*	
	Coeff.	**T Value**	**Coeff.**	**T Value**	**Coeff.**	**T Value**
LPROXIMITY	-0.0561***	-15.2100	-0.0196***	-4.0500	-0.0185***	-3.8100
NONMON			1.1116***	24.9700	1.1201***	25.1700
BIGN			3.3099***	22.2000	3.3012***	21.9400
HHI			0.7688***	7.3600	0.8546***	8.0700
MA			1.4515***	32.2300	1.4490***	31.3000
SGROWTH			0.5303***	3.8700	0.7002***	5.0500
ROA			-2.6234***	-8.9800	-2.6215***	-8.9800
Z			0.0560***	2.9800	0.0830***	4.3500
REST			1.2809***	9.2900	1.1896***	8.6100
MW			2.0042***	4.1500	1.8888***	3.9100
LNAT			0.0547***	3.0300	0.1823***	8.0400
COMPLEX			1.6402***	26.7400	1.6033***	26.0900
LIT			1.0497***	3.9900	0.9331***	3.5400
NATIONAL			-0.3532***	-15.8700	-0.3517***	-15.8100
INTERNATIONAL			0.3253***	4.5600	0.2862***	4.0100
N_DIRECTORS					-0.1007***	-7.4400
PCT_FEMALE					-1.9750***	-7.4100
F Value	176.05		198.91		198.36	
R^2	43.99%		47.87%		47.92%	
Cluster	FF48 & Year		FF48 & Year		FF48 & Year	
N	91395		91395		91395	

*, **, *** represents significance level 10%, 5% and 1% respectively. Definitions of variables are in Appendix B. Table 5 reports the results on the association between geographic proximity (LPROXIMITY) and firms' total donations (LTOTALDON). Model (1) regresses LTOTALDON on LPROXIMITY only. Model (2) adds firm-level control variables, and Model (3) further includes governance controls. All regressions are estimated with standard errors clustered by Fama–French 48 industry and year.

Table 6: Cross-Section Analyses on the Determinants of Charitable Donations

Panel A: Corporate Charitable Donations for Firms with Foundations

	Model (1) *DV=LTOTALDON*		Model (2) *DV=LTOTALDON*		Model (3) *DV=LTOTALDON*	
	Coeff.	**T Value**	**Coeff.**	**T Value**	**Coeff.**	**T Value**
LPROXIMITY	0.0287	1.44	0.1181***	6.00	0.1132***	5.74
FOUNDATION	-0.0870***	-4.43	-0.1350***	-6.97	-0.1297***	-6.68
LPROXIMITY* FOUNDATION	-2.5863***	-18.16	-2.3043	-16.52	-2.3568***	-16.72
NONMON			0.9436***	20.91	0.9398***	20.81
BIGN			3.4116***	22.94	3.3056***	22.03
HHI			0.0395	0.35	0.0081	0.07
MA			1.3968***	31.04	1.4443***	31.28
SGROWTH			0.5429***	3.97	0.6719***	4.86
ROA			-2.5903***	-8.89	-2.5363***	-8.71
Z			0.0802***	4.26	0.1023***	5.37
REST			1.1367***	8.26	1.0601***	7.69
MW			2.2045***	4.58	2.1331***	4.43
LNAT			0.0416**	2.31	0.1716***	7.59
COMPLEX			1.5302***	24.93	1.4822***	24.1
LIT			1.0348***	3.94	0.9847***	3.74
NATIONAL			-0.3463***	-15.51	-0.3447***	-15.45
INTERNATIONAL			0.3224***	4.53	0.2946***	4.13
N_DIRECTORS					-0.0607***	-4.46
PCT_FEMALE					-2.7698***	-10.33
F Value	180.64		200.27		199.83	
R^2	44.75%		48.16%		48.22%	
Cluster	FF48 & Year		FF48 & Year		FF48 & Year	
N	91395		91395		91395	

Panel B: Corporate Charitable Donations by Categories

	Model (1) *DV=LTOTALDON*		Model (2) *DV=LTOTALDON*		Model (3) *DV=LTOTALDON*	
	Coeff.	**T Value**	**Coeff.**	**T Value**	**Coeff.**	**T Value**
LPROXIMITY*EDUCATION	0.0452***	2.77	0.0424***	2.69	0.0397**	2.52
LPROXIMITY*RELIGION	0.0994***	4.1	0.0962***	4.12	0.0967***	4.14
LPROXIMIT*CIVICPUBLICAFFAIRS	0.0970***	5.78	0.0737***	4.55	0.0743***	4.58
LPROXIMITY*SCIENCE	-0.0718	-0.39	-0.0514	-0.29	-0.0557	-0.32
LPROXIMITY*HEALTH	0.1356***	7.5	0.0966***	5.52	0.0979***	5.6
LPROXIMIT*ARTS_HUMAN	-0.0140***	-14.91	-0.0138***	-15.05	-0.0137***	-14.92
LPROXIMIT*ENVIRONMEN	0.0822*	1.92	0.0175	0.42	0.0220	0.53
LPROXIMIT*SOCIALSERVICE	0.0646***	3.96	0.0383**	2.43	0.0401**	2.55
LPROXIMITY	-0.0968***	-6.22	-0.0448***	-2.92	-0.0439***	-2.86
EDUCATION	0.0734	0.82	0.0690	0.8	0.0889	1.03
RELIGION	-0.3369*	-2.43	-0.3083**	-2.3	-0.3045**	-2.27
CIVICPUBLICAFFAIRS	-0.3499***	-3.79	-0.2601***	-2.92	-0.2561***	-2.88
SCIENCE	1.0026	0.93	0.8816	0.85	0.8815	0.85

HEALTH	-0.8975***	-9.08	-0.7237***	-7.57	-0.7239***	-7.57
ARTS_HUMANITIES	0.0479***	6.98	0.0771***	11.49	0.0792***	11.78
ENVIRONMENT	0.6342**	2.18	0.3248	1.15	0.2804	0.99
SOCIALSERVICES	-0.3011***	-3.38	-0.2121**	-2.46	-0.2158**	-2.51
NONMON			1.0977***	24.42	1.1101***	24.69
BIGN			3.2920***	22.14	3.3098***	22.06
HHI			0.7015***	6.68	0.8083***	7.59
MA			1.4712***	32.75	1.4601***	31.63
SGROWTH			0.4126***	3.01	0.5798***	4.19
ROA			-2.5823***	-8.85	-2.6084***	-8.93
Z			0.0715***	3.8	0.0976***	5.13
REST			1.2855***	9.37	1.1935***	8.68
MW			1.9110***	3.97	1.8045***	3.75
LNAT			0.0515***	2.86	0.1736***	7.69
COMPLEX			1.6388***	26.81	1.6024***	26.16
LIT			1.0733***	4.09	0.9525***	3.63
NATIONAL			-0.3268***	-14.67	-0.3239***	-14.54
INTERNATIONAL			0.3386***	4.75	0.2986***	4.18
N_DIRECTORS					-0.1053***	-7.77
PCT_FEMALE					-1.7563***	-6.59
F Value	173.91		195.84		195.30	
R^2	44.65%		48.42%		48.47%	
Cluster	FF48 & Year		FF48 & Year		FF48 & Year	
N	91395		91395		91395	

*, **, *** represents significance level 10%, 5% and 1% respectively. Definitions of variables are in Appendix B. Panel A of Table 6 reports the results on the association between the interaction term of geographic proximity and the presence of a foundation (FOUNDATION*LPROXIMITY) and firms' total donations (LTOTALDON). Model (1) regresses LTOTALDON on FOUNDATION*LPROXIMITY only. Model (2) adds firm-level control variables, and Model (3) further includes governance controls. Panel B of Table 6 reports the results on the association between the interaction term of geographic proximity and various categories of donations (LPROXIMITY*EDUCATION, LPROXIMITY*RELIGION, LPROXIMIT*CIVICPUBLICAFFAIRS, LPROXIMITY*SCIENCE, LPROXIMITY*HEALTH, LPROXIMIT*ARTS_HUMAN, LPROXIMIT*ENVIRONMEN, LPROXIMIT*SOCIALSERVICE) and firms' total donations (LTOTALDON). Model (1) regresses LTOTALDON on the interaction terms only. Model (2) adds firm-level control variables, and Model (3) further includes governance controls. All regressions are estimated with standard errors clustered by Fama–French 48 industry and year.

Appendix C: Selected Figures from Preliminary Database

Figure 3: Total Donations by Type over the Years.

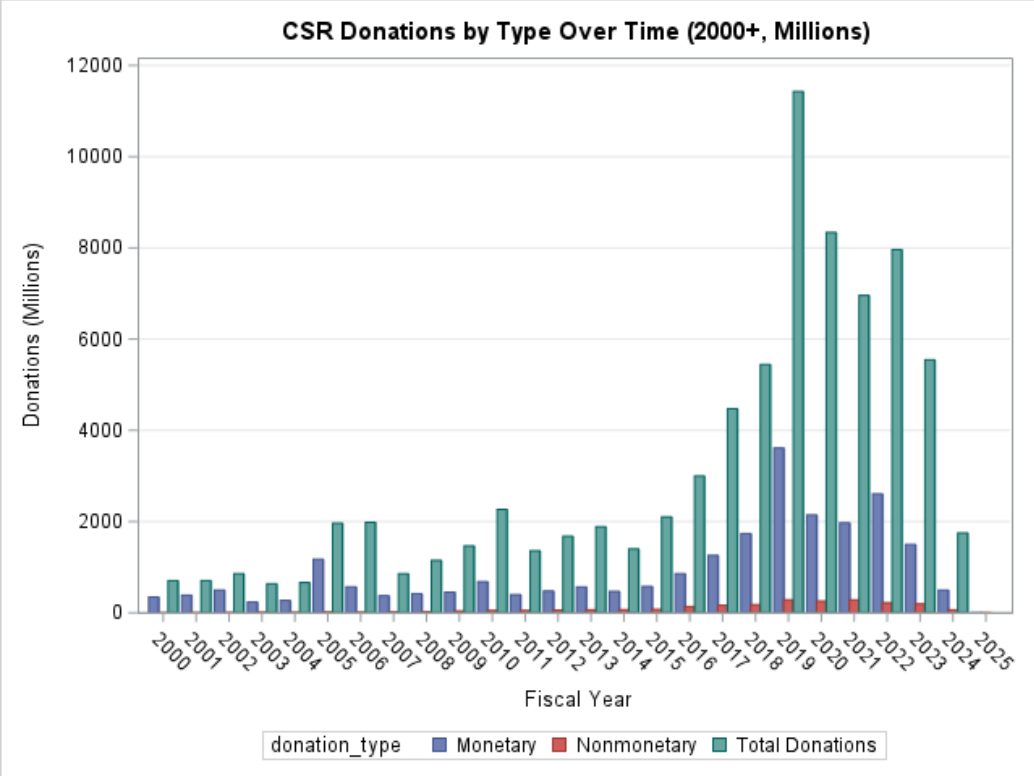


Figure 4: **Total Donations for Firms with and without Foundation.**

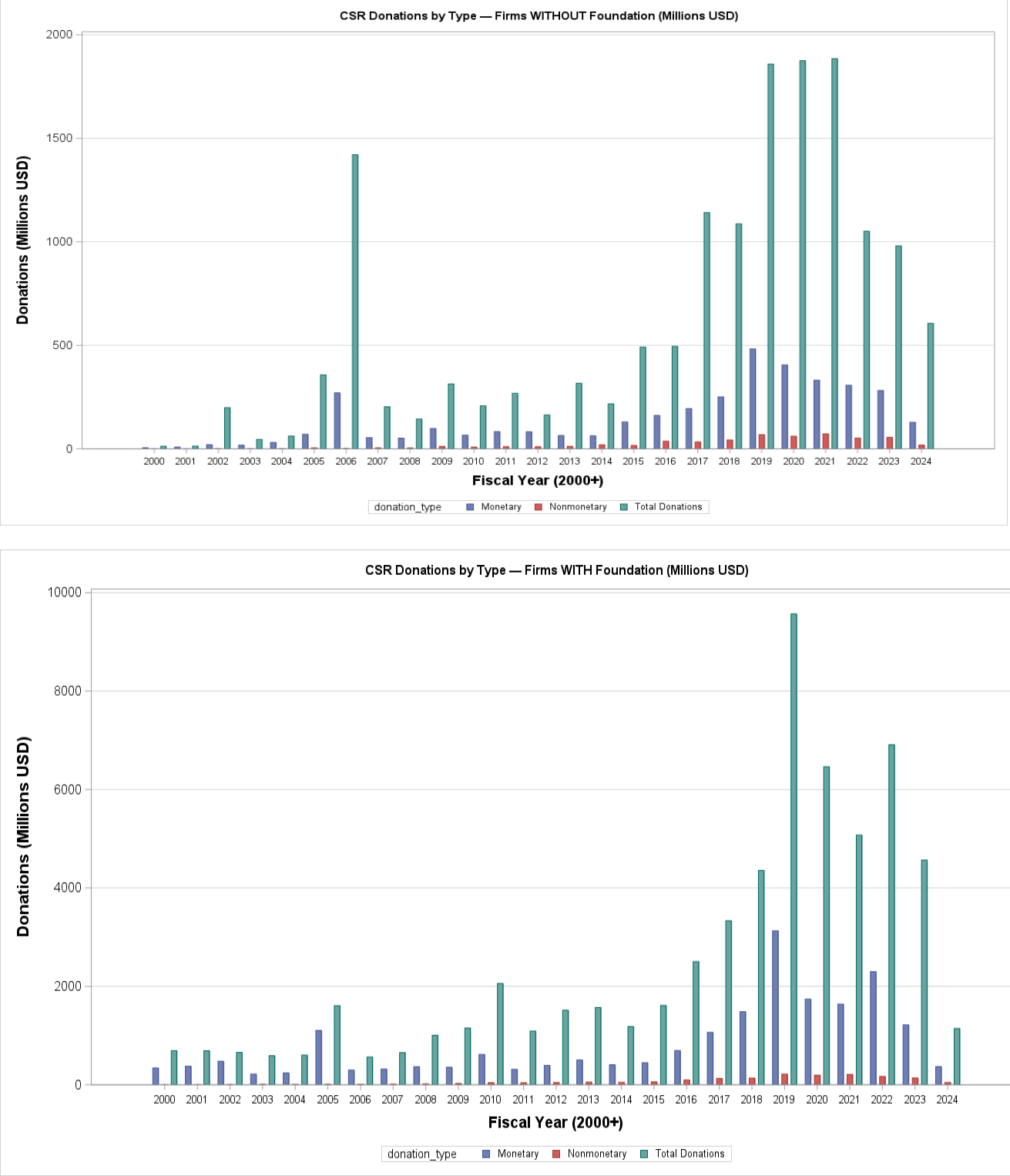


Figure 5: Total Donations by Industry.

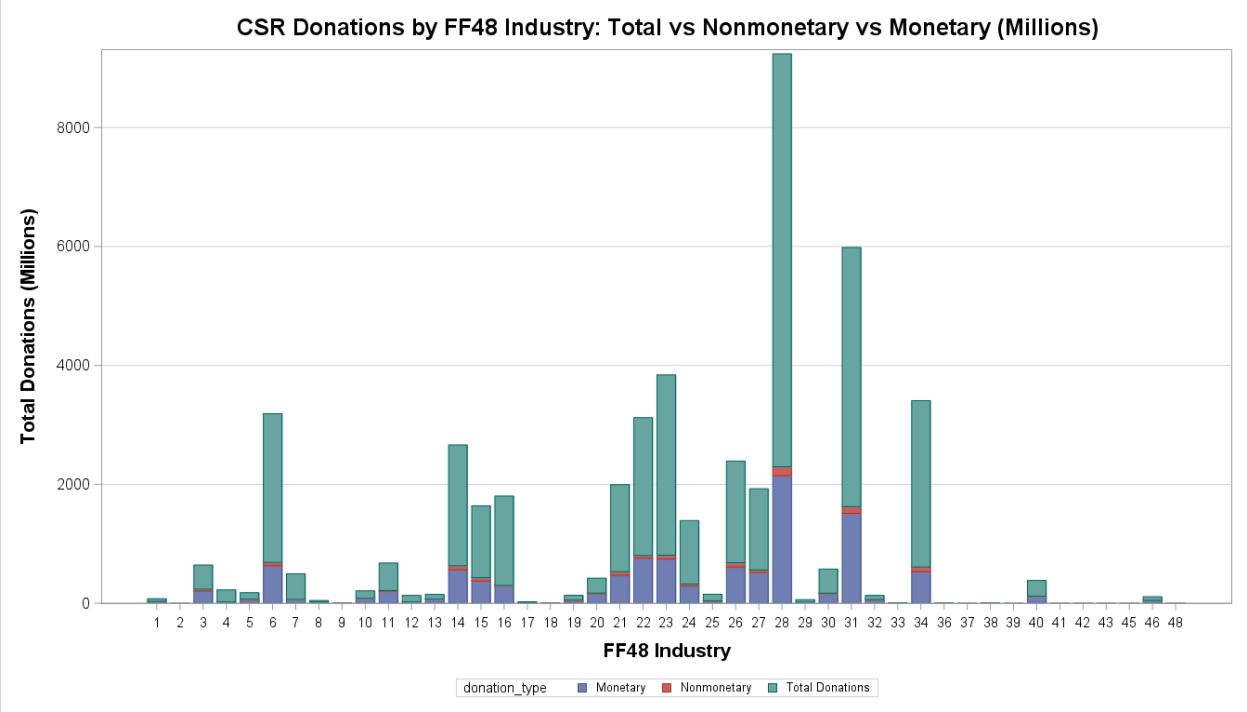


Figure 6: Average Proximity over Time.

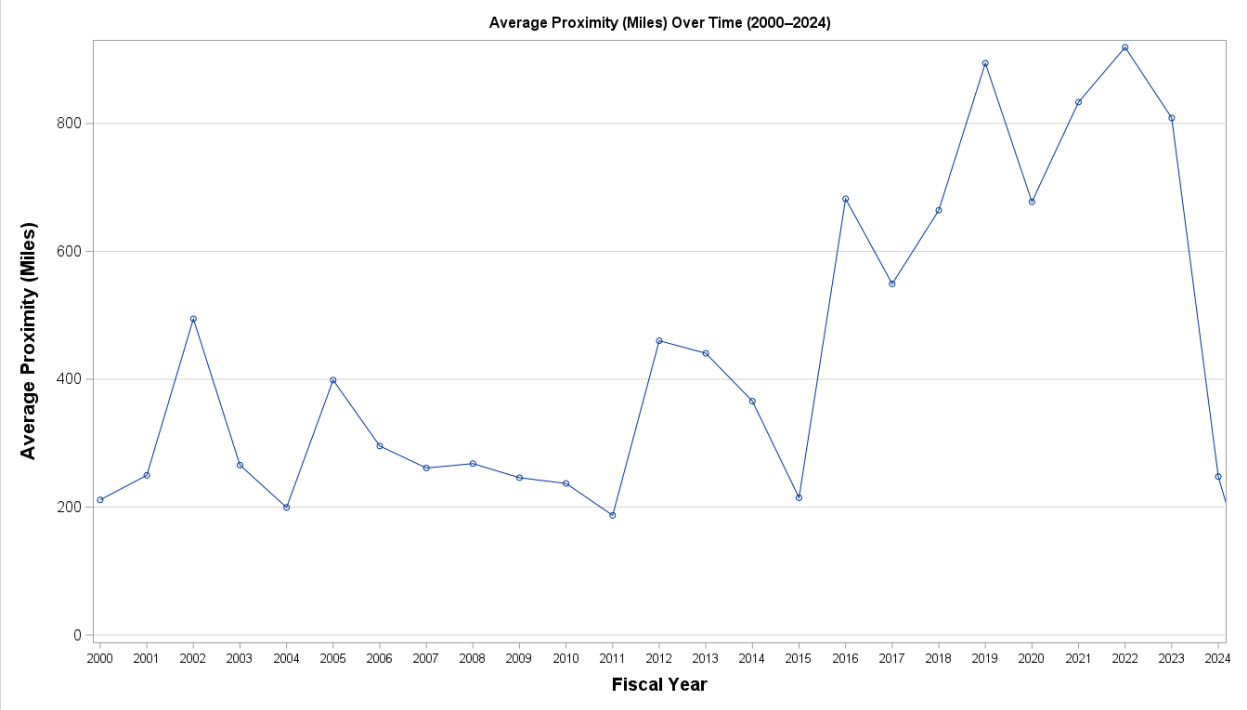
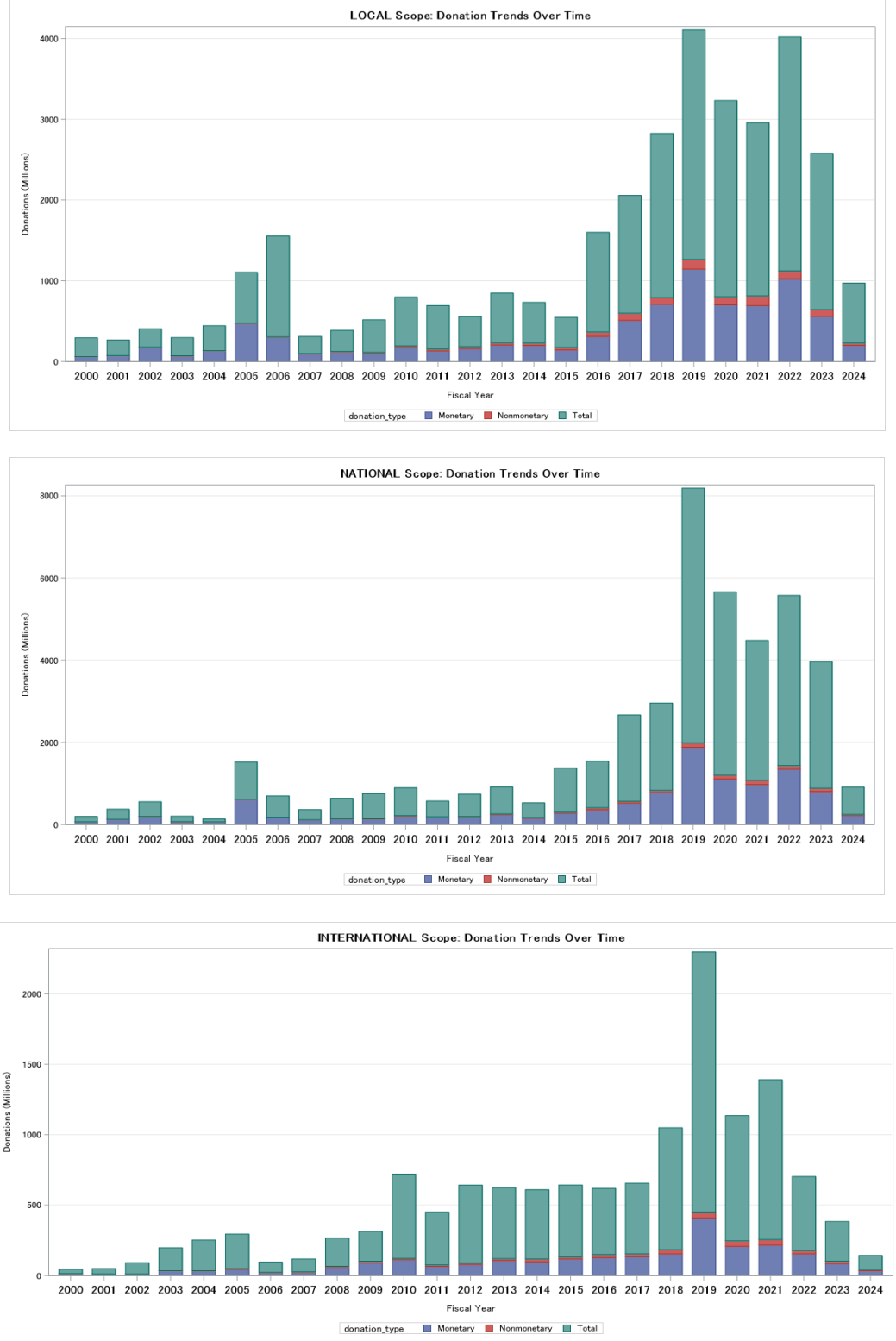


Figure 7: Donations by Scope (National, Local, and International).



References

Ali, W., Frynas, J. G., and Mahmood, Z. (2017). Determinants of corporate social responsibility (CSR) disclosure in developed and developing countries: A literature review. *Corporate social responsibility and environmental management*, *24*(4), 273-294.

Aupperle, K., Carroll, A., and Hatfield, J. (1985). An empirical examination of the relationship between corporate social responsibility and profitability. Academy of Management Journal 28 (2): 446-463. https://doi.org/10.5465/256210.

Benabou, R. and Tirole, J. (2010) Individual and corporate social responsibility. Economica 77: 1–19.

Brammer, S. and Pavelin, S. (2005). Corporate community contributions in the United Kingdom and the United States. Journal of Business Ethics 56: 15-26.

Brown, E. (1999). Assessing the value of volunteer activity. *Nonprofit and Voluntary Sector Quarterly, 28*(1), 3–17. https://doi.org/10.1177/0899764099281001

Choi, J.-H., Kim, J.-B., Qiu, A. A., and Zang, Y. (2012). Geographic Proximity between Auditor and Client: How Does It Impact Audit Quality? *Auditing: A Journal of Practice and Theory*, 31(2), 43–72. https://doi.org/10.2308/ajpt-10241

Choi, S., Mattila, A., and Quadri-Felitti, D. (2019). Donation appeals rewarding fitness in the context of CSR initiatives. Journal of Services Marketing, 33 (2):160-167.

Cornell, B., and Shapiro, A. (1987). Corporate stakeholders and corporate finance. Financial Management 16: 5-14. DOI: 10.2307/3665543.

Cracolici, M., Cuffaro, M., and Nijkamp, P. (2010). The Measurement of Economic, Social and Environmental Performance of Countries: A Novel Approach. Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement 95: 339–56.

Crifo, P. and Forget, V. (2015). The economics of corporate social responsibility: a firm-level perspective survey. Journal of Economic Surveys 29 (1): 112-130.

Deng, C., Kanagaretnam, K., and Zhou, Z. (2020). Do Locally Based Independent Directors Reduce Corporate Misconduct? Evidence from Chinese Listed Firms. *Journal of International Accounting Research*, 19(3), 61-90. https://doi.org/10.2308/jiar-19-513

Ding, D. K., Ferreira, C. and U. Wongchoti. (2019). The geography of CSR. International Review of Economics & Finance 59:265–88. doi:10.1016/j.iref.2018.09.003

European Commission. (2011). Corporate Social Responsibility: a new definition, a new agenda for action. Memo/11/730. Available at: https://ec.europa.eu/commission/presscorner/detail/en/MEMO_11_730

Ewans, M., Xiao, K., and Xu, T. (2021). Regulatory Costs of Being Public: Evidence from Bunching Estimation. NBER Working Paper 29143.

Fatima, T., and Elbanna, S. (2022). Corporate social responsibility (CSR) implementation: A review and a research agenda towards an integrative framework. Journal of Business Ethics, 183, 105–121. https://doi. org/10.1007/s10551-022-05047-8

Fennell, P. B., Folse, J. A. G., Burton, S., and Netemeyer, R. G. (2020). How donation amount and quantifier type influence consumer responses to cause-related marketing. *Journal of Business Research, 110*, 464–473. https://doi.org/10.1016/j.jbusres.2019.11.002

Gao, L., and Zhang, J. (2015). Firms' earning smoothing, corporate social responsibility, and valuation. Journal of Corporate Finance 32: 108-127. https://doi.org/10.1016/j.jcorpfin.2015.03.004

Grau, S.L. and Folse, J.A.G. (2007). Cause-related marketing (CRM): the influence of donation proximity and message-framing cues on the less-involved consumer. Journal of Advertising 36 (4):19-33.

Hemingway, C., and Maclagan, P. (2004). Managers' personal values as drivers of corporate social responsibility. Journal of Business Ethics 50: 33-44. https://doi.org/10.1023/B:BUSI.0000020964.80208.c9

Holcomb, J.L., Upchurch, R.S. and Okumus, F. (2007), "Corporate social responsibility: what are top hotel companies reporting?", International Journal of Contemporary Hospitality Management, Vol. 19 No. 6, pp. 461-475.

Huang, S.-M., Wang, T., Yen, J.-C., Lee, C.-B., Wang, Y.-C., and Yang, Y.-T. (2022). The Use of Geographic Information in Audit Data Analytics for Evidence Gathering: A Design Science Approach. *Journal of Information Systems*, 36(3), 115–128. https://doi.org/10.2308/isys-2020-045

ILO Manual on the Measurement of Volunteer Work: https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@dgreports/@stat/documents/publication/wcms_162119.pdf

Jennings, J., Lee, J., and Matsumoto, D. (2017). The Effect of Industry Co-Location on Analysts' Information costs. *Accounting Review*, 92(6), 103-127. https://doi.org/10.2308/accr-51727

Jensen MC. (2002). Value maximization, stakeholder theory, and the corporate objective function. Business Ethics Quarterly 12(2): 235–256.

Jo, H., & Harjoto, M. A. (2012). The causal effect of corporate governance on corporate social responsibility. *Journal of business ethics*, 106(1), 53-72.

Krappel, T., Bogun, A., and Borth, D. (2021). Heterogeneous Ensemble for ESG Ratings Prediction. arXiv 2021, arXiv:2109.10085.

Lai, S., Li, Z., and Yang, Y. G. (2020). East, West, Home's Best: Do Local CEOs Behave Less Myopically? *Accounting Review*, 95(2), 227–255. https://doi.org/10.2308/accr-52555

Margolis, J. and Elfenbein, H. and Walsh, J. (2009). Does it Pay to Be Good...And Does it Matter? A Meta-Analysis of the Relationship between Corporate Social and Financial Performance. Available at SSRN: https://ssrn.com/abstract=1866371 or http://dx.doi.org/10.2139/ssrn.1866371

McGuire, J., Sundgren, A., and Schneeweis, T. (1988). Corporate social responsibility and firm financial performance. The Academy of Management Journal, 31 (4): 854-872. DOI: 10.2307/256342

Pagano, M., and Volpin, P. (2005). Managers, workers, and corporate control. Journal of Finance 60 (2): 841-868.

Perez-Batres, L., Doh, J., Miller, V., and Pisani, M. (2012). Stakeholder Pressures as Determinants of CSR Strategic Choice: Why do Firms Choose Symbolic Versus Substantive Self-Regulatory Codes of Conduct? Journal of Business Ethics 110 (2): 157–172.

Peters, A. (2022). 68% of U.S. execs admit their companies are guilty of greenwashing. FastCompany. Available at: https://www.fastcompany.com/90740501/68-of-u-s-execs-admit-their-companies-are-guilty-of-greenwashing

Raghunandan, A., and Rajgopal, S. (2020). Do the socially responsible walk the talk? Working paper available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3609056.

Salamon, L. M., Sokolowski, S. W., and Haddock, M. A. (2011). Measuring the economic value of volunteer work globally: Concepts, estimates, and a roadmap to the future. *Annals of Public and Cooperative Economics*, 82(3), 217–252. https://doi.org/10.1111/j.1467-8292.2011.00437.x

Saiia, D. H, Carroll A., and Buchholtz, A. (2003). Philanthropy as Strategy: When Corporate Charity Begins at Home. Business and Society 42 (2):169–201.

Soloveichik, R. (2023). Capitalizing Data: Case Studies of Tax Forms and Individual Credit Reports. BEA Working Paper 2023-7.

Splechtna, R., Hulka, T., Sardana, D., Chandrashekar, N. D., Gracanin, D., and Matkovic, K. (2023). Interactive exploration of complex heterogeneous data: A use case on understanding city economics. In Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications — IVAPP, pages 214–221.

Teoh, T., Heng, Q., Chia, J., Shie, J., Liaw, S., Yang, M., and Nguwi, Y. (2019). Machine Learning-based Corporate Social Responsibility Prediction. In 2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM). 501–505.

Udayasankar, K. (2008). Corporate social responsibility and firm size. *Journal of business ethics*, 83(2), 167-175.

Van Marrewijk, M. (2003). Concepts and definitions of corporate social responsibility and corporate sustainability: Between agency and communion. Journal of Business Ethics 44(2-3): 95-105.